```
RESULT 15
AF254072
LOCUS       AF254072                912 bp    mRNA    linear   PLN 08-MAY-2000
DEFINITION  Zea mays histone deacetylase 2 isoform b mRNA, complete cds.
ACCESSION   AF254072
VERSION     AF254072.1  GI:7716947
KEYWORDS    .
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 912)
  AUTHORS   Dangl,M., Brosch,G., Haas,H., Loidl,P. and Lusser,A.
  TITLE     Molecular characterization of type-2 histone deacetylases in higher
            plants
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 912)
  AUTHORS   Lusser,A. and Loidl,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-APR-2000) Microbiology, Medical School, University of
            Innsbruck, Fritz-Pregl Str. 3/II, Innsbruck 6020, Austria
FEATURES             Location/Qualifiers
     source          1. .912
                     /organism="Zea mays"
                     /db_xref="taxon:4577"
     CDS             1. .912
                     /note="HD2b"
                     /codon_start=1
                     /product="histone deacetylase 2 isoform b"
                     /protein_id="AAF68624.1"
                     /db_xref="GI:7716948"
                     /translation="MEFWGLEVKPGSTVKCEPGHGFILHVSQAALGESKKSDSALMYV
                     KVDDKKLAIGTLSIDKYPQIQFDLVFNKEFELSHTSKTTSVFFSGYKVEQPIEGDEMD
                     LDSEDEEEELNIPVIKENGKADGKEEQKNQEKAVAATASKSSLGLEKKSKDDSDDSDE
                     DESDDSDEDDSDDSDEGEGLSPDEGDDDSSDEDDTSDDDEEETPTPKKPEAGKKRGAE
                     NALKTPLSDKKAKVATPPAQKTGGKKGATHVATPHPAKGKTPANNDKLTEKSPKSGGS
                     VPCKSCSKTFNSEMALQAHSKAKHGAK"
BASE COUNT      294 a    182 c    240 g    196 t
ORIGIN
```

Query Match 11.1%; Score 103.8; DB 8; Length 912;
Best Local Similarity 53.7%; Pred. No. 9.2e-13;
Matches 266; Conservative 0; Mismatches 217; Indels 12; Gaps 2;

```
Qy      49 atggagttctggggaattgaagttaaatcaggaaagccagttacagtgactcctgaagaa 108
           ||||||||||||||  | || || ||| | |||    | || |        |||| | |
Db       1 ATGGAGTTCTGGGGTCTCGAGGTCAAACCTGGATCCACTGTCAAGTGTGAGCCTGGACAT 60

Qy     109 ggcattcttatccacgtttctcaggcatcgcttggagaatgtaaaaacaagaagggagag 168
           ||| || |  | |||||||| |||||  | ||||| ||||  || || |   |  | |
Db      61 GGCTTTATCCTGCACGTTTCCCAGGCTGCCCTTGGGGAATCAAAGAAAAGTGACAGTGCC 120

Qy     169 tttgtgcctttacatgtaaaggttgggaaccagaacttggttctgggaactctatcgact 228
           ||  ||       |||| || ||||   || ||||  | |   | |||| || || |
Db     121 TTAATG------TATGTCAAAGTTGATGACAAGAAGCTTGCCATTGGAACGCTCTCTATC 174

Qy     229 gagaacatccctcagcttttctgtgatttggtattcgacaaggagtttgagctttctcac 288
           || ||   ||| ||| |    | |||||||| ||| | || ||||||||||| || |||
Db     175 GACAAATACCCACAGATACAATTCGATTTGGTTTTCAATAAAGAGTTTGAGCTGTCACAC 234

Qy     289 acttggggaaaaggaaqtgtttactttgttggatacaaaact------cccaacattgag 342
           || | |  ||     |||| |  ||   ||| |||||   |      || |     |
Db     235 ACATCGAAAACTACCAGTGTATTTTTCTCTGGTTACAAGGTTGAGCAGCCAATTGAGGGA 294

Qy     343 ccacaaggctattctgaggaagaagaggaagaagaggaagaagttcctgctgggaatgct 402
               ||   ||  |||    || || ||||| |||||   |     | | |  |
Db     295 GATGAAATGGATCTTGATTCTGAGGATGAAGAGGAGGAGCTAAACATTCCAGTAATCAAG 354

Qy     403 gccaaggctgtagctaaaccaaaggctaagcctgcagaagtgaagccagctgttgatgat 462
           |  || |    |||| |    ||||   |||    | |     ||   || || | || |
```

```
Db     355 GAAAATGGCAAAGCTGATGGGAAGGAGGAGCAGAAAAATCAAGAGAAGGCAGTAGCTGCT 414

Qy     463 gaagaggatgagtctgattctgacggaatggatgaagatgattctgatggtgaggattct 522
             ||      | ||   |  || |       | ||     |   |||   ||| ||||||
Db     415 ACAGCTTCAAAATCAAGTCTTGGCCTTGAAAAGAAAAGCAAGGATGACTCTGATGATTCT 474

Qy     523 gaggaagaagagcct 537
           || || || ||| ||
Db     475 GATGAGGATGAGTCT 489
```

Search completed: May 2, 2002, 20:13:40
Job time: 15434 sec

```
RESULT 13
AF384033
LOCUS       AF384033               1124 bp    mRNA    linear   PLN 26-JUN-2001
DEFINITION  Zea mays HD2 type histone deacetylase HDA106 (hda106) mRNA,
            complete cds.
ACCESSION   AF384033
VERSION     AF384033.1  GI:14550105
KEYWORDS    .
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 1124)
  AUTHORS   Chandler,V.L., Kaeppler,S.M., Kaeppler,H.F. and Cone,K.C.
  TITLE     Sequences from the Plant Chromatin Consortium (NSF Plant Genome
            project 9975930)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1124)
  AUTHORS   Bergstrom,D., Springer,N.M., Schmitt,L.T., Guthrie,E.,
            Sidorenko,L., Selinger,D., Kaeppler,S.M. and Cone,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-MAY-2001) Biological Sciences, Univ. of Missouri, 101
            Tucker Hall, Columbia, MO 65203, USA
FEATURES             Location/Qualifiers
     source          1. .1124
                     /organism="Zea mays"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
     gene            1. .1124
                     /gene="hda106"
     CDS             120. .980
                     /gene="hda106"
                     /codon_start=1
                     /product="HD2 type histone deacetylase HDA106"
                     /protein_id="AAK67143.1"
                     /db_xref="GI:14550106"
                     /translation="MEVGGQEVKPGATVSCKVGDGLVIHLSQAALGESKKASENAILS
                     VNIDDKKLVLGTLSVEKHPQISCDLVFDKDFELPHNSKTRSVFFRGYKSPVPLFESNS
                     GEDSSDEELKTDQIPLQNNEIKISAAKVPAKDDDDDVFIILAMMMMIYSSDDDDDDFT
                     TSDSDNEMSEEDDSSDEDEMSEEDDSSDEDEMSGGADPSDDSSDESGSEHTSAPKKTD
                     VVVGKKRAIKAEAPYGKKAKSEQSSQKTGDKASTSHPAKQSIKTPADKSRKTPTADKK
                     SPKSGSHGCK"
BASE COUNT      327 a    237 c    265 g    295 t
ORIGIN
```

Query Match 11.6%; Score 109.2; DB 8; Length 1124;
Best Local Similarity 55.2%; Pred. No. 5.8e-14;
Matches 235; Conservative 0; Mismatches 188; Indels 3; Gaps 1;

```
Qy      41 tcttagccatggagttctggggaattgaagttaaatcaggaaagccagttacagtgactc 100
           ||   || ||||||| || ||||    ||||| ||  |||||     | ||| |    |
Db     112 TCGGCGCAATGGAGGTCGGGGGTCAAGAAGTGAAGCCAGGAGCCACGGTTTCTTGCAAAG 171

Qy     101 ctgaagaaggcattcttatccacgtttctcaggcatcgcttggagaatgtaaaaacaaga 160
           ||  || ||| |  |||||||| |||| |||||  | || || ||||    ||| ||
Db     172 TTGGTGATGGCTTGGTTATCCACCTTTCACAGGCTGCACTGGGGGAATC---AAAGAAAG 228

Qy     161 agggagagtttgtgcctttacatgtaaaggttgggaaccagaacttggttctgggaactc 220
           | | |||  ||    |||  ||| ||  |||   |  ||||  | || || |||||
Db     229 CGAGTGAGAATGCCATTTTGTCTGTTAATATTGATGATAAGAAACTAGTGCTTGGAACCT 288

Qy     221 tatcgactgagaacatccctcagcttttctgtgatttggtattcgacaaggagtttgagc 280
           ||||   ||||||    |||||| | |  |||||| | ||||| || ||||| |||||||
Db     289 TATCAGTTGAGAAGCATCCTCAGATCTCTTGTGATCTTGTATTTGATAAGGATTTTGAGC 348

Qy     281 tttctcacacttggggaaaaggaagtgtttacttttgttggatacaaaactcccaacattg 340
           |  | |||| ||    || | | || |||| |||   ||| |||||  | ||      |
Db     349 TACCACACAATTCAAAAACACGTAGCGTTTTCTTCCGTGGCTACAAGTCCCCTGTTCCTC 408

Qy     341 agccacaaggctattctgaggaagaagaggaagaagaggaagaagttcctgctgggaatg 400
```

```
           |   |   ||||||  |||||       || || ||   |    | | |   | |
Db     409 TGTTTGAGTCTAATTCTGGTGAAGATAGTTCCGATGAAGAGCTAAAAACCGATCAAATTC 468

Qy     401 ctgccaaggctgtagctaaaccaaaggctaagcctgcagaagtgaagccagctgttgatg 460
           |    || |   | |    ||   ||  | |     |       |  | || ||||
Db     469 CATTGCAGAATAATGAAATTAAAATTTCTGCGGCAAAGGTTCCTGCGAAGGATGATGATG 528

Qy     461 atgaag 466
           |||| |
Db     529 ATGATG 534
```

```
RESULT 14
AR168370
LOCUS       AR168370               1245 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 15 from patent US 6287843.
ACCESSION   AR168370
VERSION     AR168370.1  GI:17904250
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1245)
  AUTHORS   Baldwin,D.Adelphi., Briggs,S.P. and Crane,V.C.
  TITLE     Maize histone deacetylases and their use
  JOURNAL   Patent: US 6287843-A 15 11-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .1245
                     /organism="unknown"
BASE COUNT      385 a    250 c    324 g    286 t
ORIGIN
```

Query Match 11.1%; Score 104.2; DB 6; Length 1245;
Best Local Similarity 53.7%; Pred. No. 7.2e-13;
Matches 267; Conservative 0; Mismatches 218; Indels 12; Gaps 2;

```
Qy      47 ccatggagttctggggaattgaagttaaatcaggaaagccagttacagtgactcctgaaq 106
           | ||||||||||||||||  | || || ||| | |||     | || |         |||| |
Db      82 CGATGGAGTTCTGGGGTCTCGAGGTCAAACCTGGATCCACTGTCAAGTGTGAGCCTGGAC 141

Qy     107 aaggcattcttatccacgtttctcaggcatcgcttggagaatgtaaaaacaagaaggggag 166
           | ||| || |  | |||||||| |||||  | ||||| |||  || || |   | | |
Db     142 ATGGCTTTATCCTGCACGTTTCCCAGGCTGCCCTTGGGGAATCAAAGAAAAGTGACAGTG 201

Qy     167 agtttgtgcctttacatgtaaaggttgggaaccagaacttggttctgggaactctatcga 226
           ||  ||       |||| || ||||   || ||||  | |   | ||||| || || |
Db     202 CCTTAATG------TATGTCAAAGTTGATGACAAGAAGCTTGCCATTGGAACGCTCTCTA 255

Qy     227 ctgagaacatccctcagcttttctgtgatttggtattcgacaaggagtttgagctttctc 286
           ||  ||   ||| ||| |    |  |||||||| ||| | || ||||||||||| || |
Db     256 TCGACAAATACCCACAGATACAATTCGATTTGGTTTTCAATAAAGAGTTTGAGCTGTCAC 315

Qy     287 acacttqqqqaaaaqqaaqtgtttacttttgttggatacaaaact------cccaacattg 340
           |||| | |  ||     ||||| |  ||   ||| |||||    |      || |    |
Db     316 ACACATCGAAAACTACCAGTGTATTTTTCTCTGGTTACAAGGTTGAGCAGCCAATTGAGG 375

Qy     341 agccacaaggctattctgaggaagaagaggaagaagaggaagaagttcctgctgggaatg 400
           |     ||  |||    || || ||||| |||||   |     | | |  |
Db     376 GAGATGAAATGGATCTTGATTCTGAGGATGAAGAGGAGGAGCTAAACATTCCAGTAATCA 435

Qy     401 ctgccaaggctgtagctaaaccaaaggctaagcctgcagaagtgaagccagctgttgatg 460
           |  || |    |||| |    ||||   |||    | |     ||   || || | ||
Db     436 AGGAAAATGGCAAAGCTGATGGGAAGGAGGAGCACAAAAATCAAGAGAAGGCAGTAGCTG 495

Qy     461 atgaagaggatgagtctgattctgacggaatqgatgaagatgattctgatggtgaggatt 520
           |  ||      | ||   |  || |       | ||     |   |||  ||| ||||
Db     496 CTACAGCTTCAAAATCAAGTCTTGGCCTTGAAAAGAAAAGCAAGGATGACTCTGATGATT 555

Qy     521 ctgaggaagaagagcct 537
           |||| || || ||| ||
Db     556 CTGATGAGGATGAGTCT 572
```

```
FEATURES             Location/Qualifiers
     source          1. .885
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
     CDS             1. .885
                     /codon_start=1
                     /product="putative histone deacetylase HD2c"
                     /protein_id="AAF70197.1"
                     /db_xref="GI:7804963"
                     /translation="MEFWGVEVKNGKPLHLDPGLDRLVHISQVALGESKNNVTEPIQL
                     YVTVGSDKLLICTLSHEKFPQLSTEIVLERNFALSHTWKNGSVFFSGYKVDASDPEPE
                     DLIDDQLEAAGFKAAPKSAAKQVNFQLPNEDVKAKQDDDADGSEEDSSDDDDSENSGD
                     EEEEKVTAESDSEEDDSSDDEEDDSSEEETPKKPEEPKKRSAEPNSSKNPASNKKAKF
                     VTPQKTDSKKPHVHVATPHPSKQAGKNSGGGSTGETSKQQQTPKSAGAFGCKSCTRTF
                     TSEMGLQSHTKAKHSAAA"
BASE COUNT       266 a    190 c    231 g    198 t
ORIGIN
```

```
  Query Match             13.1%;  Score 123.4;  DB 8;  Length 885;
  Best Local Similarity   51.2%;  Pred. No. 4.4e-17;
  Matches  289;  Conservative    0;  Mismatches  276;  Indels    0;  Gaps    0;

Qy      49 atggagttctggggaattgaagttaaatcaggaaagccagttacagtgactcctgaagaa 108
           ||||||||||||||  ||||||||||    || || ||| ||    |   |||||    |
Db       1 ATGGAGTTCTGGGGTGTTGAAGTTAAGAATGGCAAACCACTTCACCTAGATCCTGGGCTA 60

Qy     109 ggcattcttatccacgtttctcaggcatcgcttggagaatgtaaaaacaagaagggagag 168
           | ||  ||  | ||| | |||||||   |  | |||||  |||| ||||        ||
Db      61 GACAGGCTCGTGCACATCTCTCAGGTTGCTTTGGGAGAGAGTAAGAACAATGTAACCGAA 120

Qy     169 tttgtgcctttacatgtaaaggttgggaaccagaacttggttctgggaactctatcgact 228
             | |   |  | ||  | |||||     | ||  |  |  | ||||| |||||   |
Db     121 CCAATCCAGCTTTACGTGACTGTTGGATCGGATAAGCTTCTCATTGGAACGCTATCTCAT 180

Qy     229 gagaacatccctcagctttctgtgatttggtattcgacaaggagtttgagctttctcac 288
           |||||  | ||||||| |     ||  | || || || | | | |||| ||| |||||
Db     181 GAGAAGTTTCCTCAGCTGTCTACGGAGATTGTCTTGGAAAGGAACTTTGCGCTGTCTCAT 240

Qy     289 acttggggaaaaggaagtgtttactttgttggatacaaaactcccaacattgagccacaa 348
           ||||||   || || || |||| |||   ||| ||||||  |         ||| ||  |
Db     241 ACTTGGAAGAATGGGAGCGTTTTCTTCTCTGGTTACAAAGTTGATGCATCTGATCCCGAG 300

Qy     349 ggctattctgaggaagaagaggaagaagaggaagaagttcctgctgggaatgctgccaag 408
             |  |  | || ||  ||   |||| |  | |       |    | |
Db     301 CCTGAGGATTTGATTGATGATCAACTTGAGGCTGCTGGTTTCAAAGCTGCTCCGAAATCT 360

Qy     409 gctgtagctaaaccaaaggctaagcctgcagaagtgaagccagctgttgatgatgaagag 468
           ||||    |   ||  | ||   || | |   |       |   |  | || ||
Db     361 GCTGCGAAACAGGTGAACTTTCAGTTGCCAAATGAAGATGTCAAAGCCAAGCAAGATGAT 420

Qy     469 gatgagtctgattctgacggaatggatgaagatgattctgatggtgaggattctgaggaa 528
           || |     | |  ||| | |     |  |||||||  |||| ||||  | ||||  ||
Db     421 GACGCTGACGGTAGTGAAGAAGATTCTTCAGATGATGATGATAGTGAAAACTCTGGAGAC 480

Qy     529 gaagagcctacacctaagaagcctgcatcaagcaagaagagagctaatgaaactaccccc 588
           || |||              |  ||| |||   ||  ||  || | |   | |   |
Db     481 GAGGAGGAAGAGAAGGTTACTGCTGAATCTGACAGTGAGGAAGATGACTCATCAGACGAC 540

Qy     589 aaagcacctgtgtcagcaaagaagg 613
             ||| |  ||  ||  || |  |||
Db     541 GAAGAAGATGACTCCTCAGAAGAGG 565
```

```
RESULT  12
AF255711
LOCUS       AF255711                 894 bp    mRNA    linear   PLN 15-MAY-2000
DEFINITION  Oryza sativa putative histone deacetylase HD2 mRNA, complete cds.
ACCESSION   AF255711
VERSION     AF255711.1  GI:7804959
KEYWORDS    .
SOURCE      Oryza sativa.
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 894)
  AUTHORS   Dangl,M., Brosch,G., Haas,H., Loidl,P., Graessle,S. and Lusser,A.
  TITLE     Molecular characterization of type-2 histone deacetylases in higher
            plants
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 894)
  AUTHORS   Dangl,M., Haas,H., Loidl,P. and Lusser,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-APR-2000) Microbiology, University of Innsbruck,
            Fritz-Pregl-Str. 3, Innsbruck 6020, Austria
FEATURES             Location/Qualifiers
     source          1. .894
                     /organism="Oryza sativa"
                     /db_xref="dbEST:D15380"
                     /db_xref="taxon:4530"
     CDS             1. .894
                     /codon_start=1
                     /product="putative histone deacetylase HD2"
                     /protein_id="AAF70196.1"
                     /db_xref="GI:7804960"
                     /translation="MEFWGLEVKPGQTVKCEPEDERFLHLSQAALGESKKGSDNAVMY
                     VKTDDQKLVIGTLSADKFPQIQFDLVFDKEFELSHTSKTASVFFSGYKVSQPAEEDEM
                     DFDSEEVEDEEEEEKIIPAPRANGKVEGKENEQKKQGKTDSSASKSKAAVNDDDDDDD
                     SAEDDSQDEDLSPEDDDDDSSEDDSSEDDEDESDEEDTPKKPETGKRKVAEIVLKTPS
                     SDKKAKIATPSGQKTGDKKGVHVATPHPAKQASKTPVNDKSKEKSPKSGGGSISCKSC
                     SKTFNSEMALQSHSKAKHPAK"
BASE COUNT       289 a    172 c    234 g    199 t
ORIGIN
```

```
  Query Match             12.5%;  Score 117.6;  DB 8;  Length 894;
  Best Local Similarity   55.3%;  Pred. No. 8.4e-16;
  Matches  272;  Conservative    0;  Mismatches  214;  Indels    6;  Gaps    2;

Qy      49 atggagttctggggaattgaagttaaatcaggaaagccagttacagtgactcctgaagaa 108
           ||||||||||||||  |||||||  ||  | ||| || | || | |      ||||||||
Db       1 ATGGAGTTCTGGGGTCTTGAAGTCAAGCCTGGACAGACTGTCAAATGTGAGCCTGAAGAT 60

Qy     109 ggcattcttatccacgtttctcaggcatcgcttggagaatgtaaaaacaagaagggagag 168
           |      || | ||| ||||||||||  | |||| ||||    |||  ||        ||
Db      61 GAACGCTTTTTGCACCTTTCTCAGGCTGCTCTTGGGGAAT---CAAAGAAAGGATCTGAC 117

Qy     169 tttgtgcctttacatgtaaaggttgggaaccagaacttggttctgggaactctatcgact 228
             ||      | |||| ||   ||   | || ||  | ||  | ||||| || ||  ||
Db     118 AATGCAGTAATGTATGTTAAAACTGATGATCAAAAGCTAGTCATTGGAACCCTCTCAGCT 177

Qy     229 gagaacatccctcagcttttctgtgatttggtattcgacaagggagtttgagctttctcac 288
           || ||  |||||||  |    | |||||||||| || ||||| ||||||||||| || |||
Db     178 GACAAGTTCCCTCAAATCCAGTTTGATTTGGTCTTTGACAAAGAGTTTGAGCTGTCACAC 237

Qy     289 acttggggaaaaggaagtgtttactttgttggatacaaaactcccaacattgagccacaa 348
           ||||      |  |  ||||| | ||||  ||| ||||||  | || |     |      ||
Db     238 ACTTCAAAGACTGCTAGTGTGTTCTTTTCTGGCTACAAAGTTTCCCAGCCGGCTGAGGAA 297

Qy     349 ggctattctgaggaagaagaggaagaagaggaagaagttcctgctgggaatgctgccaag 408
           |    |    ||    ||    |||||||  ||||| |      |  |    |       ||
Db     298 GATGAAATGGATTTTGATTCTGAAGAAGTTGAAGATGAAGAGGAGGAAGAAAAGATCATT 357

Qy     409 gctgtagctaaaccaaaggctaagcctgcagaagtgaagccagctgttgatgatgaagag 468
             | |   | |   |||| |  ||    |  |||| ||||  |  ||   |  |  || |
Db     358 CCAGCTCCCAGGGCAAATGGCAA---AGTTGAAGGGAAGGAAAATGAGCAGAAAAAACAA 414

Qy     469 gatgagtctgattctgacggaatggatgaagatgattctgatggtgaggattctgaggaa 528
           |   || | ||||||    |       |  | | | | |    ||| |||  ||| ||
Db     415 GGCAAGACAGATTCTTCAGCTTCAAAATCAAAGGCTGCAGTGAATGACGATGATGATGAT 474

Qy     529 gaagagcctaca 540
           || ||   | ||
Db     475 GATGACAGTGCA 486
```

```
                     /db_xref="GI:13926268"
                     /translation="MEFWGVEVKNGKPLHLDPGLDRLVHISQVALGESKNNVTEPIQL
                     YVTVGSDKLLIGTLSHEKFPQLSTEIVLERNFALSHTWKNGSVFFSGYKVDASDPEPE
                     DLIDDQLEAAGFKAAPKSAAKQVNFQLPNEDVKAKQDDDADGSEEDSSDDDDSENSGD
                     EEEEKVTAESDSEEDDSSDDEEDDSSEEETPKKPEEPKKRSAEPNSSKNPASNKKAKF
                     VTPQKTDSKKPHVHVATPHPSKQAGKNSGGGSTGETSKQQQTPKSAGAFGCKSCTRTF
                     TSEMGLQSHTKAKHSAAA"
     3'UTR           998. .1236
BASE COUNT      351 a    268 c    293 g    324 t
ORIGIN


  Query Match             13.2%;  Score 124.2;  DB 8;  Length 1236;
  Best Local Similarity   51.1%;  Pred. No. 2.8e-17;
  Matches  291;  Conservative    0;  Mismatches  278;  Indels    0;  Gaps    0;

Qy      45 agccatggagttctggggaattgaagttaaatcaggaaagccagttacagtgactcctga 104
           | | ||||||||||||||||  ||||||||||     || || ||| ||    |   |||||
Db     109 AACAATGGAGTTCTGGGGTGTTGAAGTTAAGAATGGCAAACCACTTCACCTAGATCCTGG 168

Qy     105 agaaggcattcttatccacgtttctcaggcatcgcttggagaatgtaaaaacaagaaggg 164
             || ||  ||  | ||| | |||||||   |  | |||||  |||| |||||
Db     169 GCTAGACAGGCTCGTGCACATCTCTCAGGTTGCTTTGGGAGAGAGTAAGAACAATGTAAC 228

Qy     165 agagtttgtgcctttacatgtaaaggttgggaaccagaacttggttctgggaactctatc 224
             ||     | |   | | || |  |||||     | ||  |  |  | |||||  |||||
Db     229 CGAACCAATCCAGCTTTACGTGACTGTTGGATCGGATAAGCTTCTCATTGGAACGCTATC 288

Qy     225 gactgagaacatccctcagcttttctgtgatttggtattcgacaaggagtttgagctttc 284
             ||||||  | |||||||| |     ||  | || || || | | | |||| ||| ||
Db     289 TCATGAGAAGTTTCCTCAGCTGTCTACGGAGATTGTCTTGGAAAGGAACTTTGCGCTGTC 348

Qy     285 tcacacttggggaaaaggaagtgtttactttgttggatacaaaactcccaacattgagcc 344
           ||| ||||||   || || || |||| |||   ||| ||||||  |         ||| ||
Db     349 TCATACTTGGAAGAATGGGAGCGTTTTCTTCTCTGGTTACAAAGTTGATGCATCTGATCC 408

Qy     345 acaaggctattctgaggaagaagaggaagaagaggaagaagttcctgctgggaatgctgc 404
             |    |   |  || ||  ||   ||||  |  | |      |    | |
Db     409 CGAGCCTGAGGATTTGATTGATGATCAACTTGAGGCTGCTGGTTTCAAAGCTGCTCCGAA 468

Qy     405 caaggctgtagctaaaccaaaggctaagcctgcagaagtgaagccagctgttgatgatga 464
             ||||      |    ||   | ||    || | |   |       |   |  | ||
Db     469 ATCTGCTGCGAAACAGGTGAACTTTCAGTTGCCAAATGAAGATGTCAAAGCCAAGCAAGA 528

Qy     465 agaggatgagtctgattctgacggaatggatgaagatgattctgatggtgaggattctga 524
             || || |     | |  ||| | |     |  |||||||  |||| ||||  | ||||
Db     529 TGATGACGCTGACGGTAGTGAAGAAGATTCTTCAGATGATGATGATAGTGAAAACTCTGG 588

Qy     525 ggaagaagagcctacacctaagaagcctgcatcaagcaagaagagagctaatgaaactac 584
             || || |||          |   ||| |||   ||   ||  || | |   | |
Db     589 AGACGAGGAGGAAGAGAAGGTTACTGCTGAATCTGACAGTGAGGAAGATGACTCATCAGA 648

Qy     585 ccctaaagcacctgtgtcagcaaagaagg 613
           |    ||| |  ||  ||  || |  |||
Db     649 CGACGAAGAAGATGACTCCTCAGAAGAGG 677


RESULT  10
AR168371
LOCUS       AR168371                1307 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 17 from patent US 6287843.
ACCESSION   AR168371
VERSION     AR168371.1  GI:17904252
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1307)
  AUTHORS   Baldwin,D.Adelphi., Briggs,S.P. and Crane,V.C.
  TITLE     Maize histone deacetylases and their use
  JOURNAL   Patent: US 6287843-A 17 11-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .1307
                     /organism="unknown"
BASE COUNT      386 a    260 c    308 g    353 t
ORIGIN


  Query Match             13.2%;  Score 123.8;  DB 6;  Length 1307;
  Best Local Similarity   54.5%;  Pred. No. 3.4e-17;
  Matches  298;  Conservative    0;  Mismatches  237;  Indels   12;  Gaps    2;

Qy      41 tcttagccatggagttctggggaattgaagttaaatcaggaaagccagttacagtgactc 100
           || | || ||||||||||||||||    ||||| ||  |||||    | ||| |    |
Db      79 TCGTCGCAATGGAGTTCTGGGGTGAAGAAGTGAAGCCAGGAGCCACGGTTTCTTGCAAAG 138

Qy     101 ctgaagaaggcattcttatccacgtttctcaggcatcgcttggagaatgtaaaaacaaga 160
             ||  || ||  |  |||||||| |||| |||||  | || || |||     ||| ||
Db     139 TTGGTGATGGTTTGGTTATCCACCTTTCACAGGCTGCCCTAGGGGAA---CCAAAGAAAG 195

Qy     161 agggagagtttgtgcctttacatgtaaaggttgggaaccagaacttggttctgggaactc 220
             | | |||  ||    | |   ||| ||  |||   |  |||| | || || |||||
Db     196 CGAGTGAGAATGCCATTGTGTCTGTCAAAATTGATGATAAGAAACTAGTGCTTGGAACCT 255

Qy     221 tatcgactgagaacatccctcagcttttctgtgatttggtattcgacaaggagtttgagc 280
           ||||  ||||||    |||||  | |  |||||| |||||| || || || ||||||
Db     256 TATCAGTTGAGAAGCATCCTCAAATCTCTTGTGATCTGGTATTTGATAAAGATTTTGAGT 315

Qy     281 tttctcacacttggggaaaaggaagtgtttactttgttggatacaaaactcccaacattg 340
           | || |||| ||     | ||  |||||||  |||   ||| |||||  | ||      |
Db     316 TATCACACAATTCAAAGACAGCTAGTGTTTTCTTCTGTGGCTACAAGTCACCTGTTCCTC 375

Qy     341 agccacaaggctattctgaggaagaagaggaagaagaggaagaagttcctgctgggaatg 400
           |    |    ||||||  |||||      ||| || || |  |  || | |  | ||
Db     376 TGTTTGAGTCTGATTCTGGTGAAGATAGTTCAGATGAAGAGGTTGAACCCGATCTAATTC 435

Qy     401 ctgccaaggctgtagctaaaccaaaggcta---------agcctgcagaagtgaagccag 451
           |    ||  |   |  |   ||  |||                || |   | ||     |
Db     436 CAATGCAGAATAATGAAATTAAAATTTCTACTGCAAAGGTTCCCGTGAAGGTTGGTATAC 495

Qy     452 ctgttgatgatgaagaggatgagtctgattctgacggaatggatgaagatgattctgatg 511
             || |||||||||  ||     |||  |  |||  |     |||      ||||  ||||
Db     496 AAAATGCTGATGAAGATGAAACTTCTAGTGGTGATGATGATGATTTCACTGATAGTGATA 555

Qy     512 gtgaggattctgaggaagaagagcctacacctaagaagcctgcatcaagcaagaagagag 571
           ||||    |||||||||||| ||| | |    | |  |    |  |||||| | |
Db     556 GTGAAATGTCTGAGGAAGATGAGTCCAGTGATGAAGATGAAGTGTCAAGCGATACAGATA 615

Qy     572 ctaatga 578
           ||| |||
Db     616 CTAGTGA 622


RESULT  11
AF255712
LOCUS       AF255712                885 bp    mRNA    linear   PLN 15-MAY-2000
DEFINITION  Arabidopsis thaliana putative histone deacetylase HD2c mRNA,
            complete cds.
ACCESSION   AF255712
VERSION     AF255712.1  GI:7804962
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 885)
  AUTHORS   Dangl,M., Brosch,G., Haas,H., Loidl,P., Graessle,S. and Lusser,A.
  TITLE     Molecular characterization of type-2 histone deacetylases in higher
            plants
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 885)
  AUTHORS   Dangl,M., Haas,H., Loidl,P. and Lusser,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-APR-2000) Microbiology, University of Innsbruck,
            Fritz-Pregl-Str. 3, Innsbruck 6020, Austria
```

```
LOCUS       AF044914                 918 bp    mRNA    linear   PLN 11-FEB-1998
DEFINITION  Arabidopsis thaliana putative histone deacetylase (HD2) mRNA,
            complete cds.
ACCESSION   AF044914
VERSION     AF044914.1  GI:2854069
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 918)
  AUTHORS   Dangl,M., Haas,H. and Loidl,P.
  TITLE     Arabidopsis thaliana complete cDNA-sequence homologous to Zea mays
            HD2
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 918)
  AUTHORS   Dangl,M., Haas,H. and Loidl,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JAN-1998) Microbiology, Medical School, University of
            Innsbruck, Fritz-Pregl-Str. 3, Innsbruck 6020, Austria
FEATURES             Location/Qualifiers
     source          1. .918
                     /organism="Arabidopsis thaliana"
                     /cultivar="Columbia"
                     /db_xref="taxon:3702"
                     /clone="133F15T7"
                     /clone_lib="EST-clone 133F15 (T. Newman, Plant Physiol.
                     106, 1241-1255, 1994)"
     gene            1. .918
                     /gene="HD2"
     CDS             1. .918
                     /gene="HD2"
                     /codon_start=1
                     /product="putative histone deacetylase"
                     /protein_id="AAC02539.1"
                     /db_xref="GI:2854070"
                     /translation="MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLS
                     VTVGGAKLVIGTLSQDKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFT
                     SSDDEDVPEAVPAPAPTAVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDES
                     DDEDESEEDDDSEKGMDVDEDDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKT
                     PVSGKKAKPAAAPASTPQKTEKKKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGGN
                     NNKKPFNSGKQFGGSNNKGSNKGKGKGRA"
BASE COUNT      292 a    194 c    239 g    193 t
ORIGIN
```

```
  Query Match           16.0%;  Score 149.8;  DB 8;  Length 918;
  Best Local Similarity  64.8%;  Pred. No. 6.6e-23;
  Matches  239;  Conservative    0;  Mismatches  127;  Indels    3;  Gaps    1;

Qy      49 atggagttctggggaattgaagttaaatcaggaaagccagttacagtgactcctgaagaa 108
           ||||||||||||||| |||  || | | ||  |||  |   ||  ||||||||||||||||
Db       1 ATGGAGTTCTGGGGAGTTGCGGTGACACCAAAAAACGCTACTAAGGTGACTCCTGAAGAA 60

Qy     109 ggcattcttatccacgtttctcaggcatcgcttggagaatgtaaaaacaagaagggagag 168
           | ||  ||| ||||| |||||||||| || |||   || || | |   ||    |||||
Db      61 GACAGCCTTGTCCACATTTCTCAGGCTTCACTT---GACTGCACAGTGAAATCTGGAGAA 117

Qy     169 tttgtgcctttacatgtaaaggttgggaaccagaacttggttctgggaactctatcgact 228
           | ||||  |||   ||| |  |||||       ||  | ||| | ||||| || ||
Db     118 TCTGTGGTTTTGAGTGTGACTGTTGGTGGGGCTAAACTTGTTATTGGAACACTTTCACAA 177

Qy     229 gagaacatccctcagcttttctgtgatttggtattcgacaaggagtttgagctttctcac 288
           || ||  |||||||| ||  || ||||||||| || || || ||||||||||||| |||
Db     178 GACAAGTTCCCTCAGATTAGCTTTGATTTGGTTTTTGATAAAGAGTTTGAGCTTTCACAC 237

Qy     289 acttggggaaaaggaagtgtttacttttgttggatacaaaactcccaacattgagccacaa 348
           |   |    |||| || |||| | ||  |||| |||||| | ||||||||  ||||   |
Db     238 AGCGGTACCAAAGCAAATGTTCATTTCATTGGCTACAAATCCCCCAACATCGAGCAGGAT 297

Qy     349 ggctattctgaggaagaagaggaagaagaggaagaagttcctgctgggaatgctgccaag 408
           | ||   ||      || || || || |     |||| |  | |||    |||  | |
```

```
Db     298 GACTTCACTAGTTCGGATGATGAGGATGTTCCTGAAGCTGTTCCTGCTCCTGCCCCTACT 357

Qy     409 gctgtagct 417
           ||||| ||
Db     358 GCTGTTACT 366
```

```
RESULT   9
AF372889
LOCUS       AF372889                 1236 bp    mRNA    linear   PLN 23-MAY-2001
DEFINITION  Arabidopsis thaliana AT5g03740/F17C15_160 mRNA, complete cds.
ACCESSION   AF372889
VERSION     AF372889.1  GI:13926267
KEYWORDS    FLI_CDNA.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1236)
  AUTHORS   Cheuk,R., Chen,H., Kim,C.J., Meyers,M.C., Shinn,P., Banh,J.,
            Bowser,L., Carninci,P., Chung,M.K., Goldsmith,A.D., Hayashizaki,Y.,
            Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
            Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Narusaka,M.,
            Nguyen,M., Palm,C.J., Pham,P.K., Quach,H.L., Sakano,H., Sakurai,T.,
            Satou,M., Seki,M., Southwick,A., Toriumi,M., Yamada,K., Yu,G.,
            Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.
  TITLE     Arabidopsis cDNA clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1236)
  AUTHORS   Cheuk,R., Chen,H., Kim,C.J., Meyers,M.C., Shinn,P., Banh,J.,
            Bowser,L., Carninci,P., Chung,M.K., Goldsmith,A.D., Hayashizaki,Y.,
            Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
            Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Narusaka,M.,
            Nguyen,M., Palm,C.J., Pham,P.K., Quach,H.L., Sakano,H., Sakurai,T.,
            Satou,M., Seki,M., Southwick,A., Toriumi,M., Yamada,K., Yu,G.,
            Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-APR-2001) Salk Institute Genomic Analysis Laboratory
            (SIGnAL), Plant Biology Laboratory, The Salk Institute for
            Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
            USA
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA') :  Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P.,  Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members carried out the
            sequencing and annotation of the RAFL cDNAs: Shinn,P., Chen,H.,
            Cheuk,R., Kim,C.J., Koesema,E., Meyers,M.C., Tracy,S.E., Banh,J.
            Bowser,L., Chung,M.K., Goldsmith,A.D., Jones,T., Karlin-Neumann,G.,
            Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M.,
            Palm,C.J., Pham,P.K., Quach,H.L., Sakano,H., Southwick,A.,
            Tang,C.C., Toriumi,M., Yamada,K., Yu,G., Davis,R.W., Theologis,A.,
            and Ecker,J.R.

            Shinn,P. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
            this work.  Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
            contributed equally to this work as PIs.
FEATURES             Location/Qualifiers
     source          1. .1236
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /chromosome="5"
                     /clone="RAFL03-05-D06(R12384)"
                     /note="ecotype: Columbia"
     5'UTR           1. .112
     CDS             113. .997
                     /note="histone deacetylase-like protein"
                     /codon_start=1
                     /product="AT5g03740/F17C15_160"
                     /protein_id="AAK49605.1"
```

```
Qy    385 gttcctgctgqqqaatgctgccaaggctgtagct 417
          | |  | |||    |||  | |  |||||  ||
Db    408 GCTGTTCCTGCTCCTGCCCCTACTGCTGTTACT 440


RESULT   6
AF195546
LOCUS       AF195546               1187 bp    mRNA    linear   PLN 01-NOV-2000
DEFINITION  Arabidopsis thaliana putative histone deacetylase (HD2B) mRNA,
            complete cds.
ACCESSION   AF195546
VERSION     AF195546.1  GI:11066136
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1187)
  AUTHORS   Wu,K., Tian,L., Malik,K., Brown,D. and Miki,B.
  TITLE     Functional analysis of HD2 histone deacetylase homologs in
            Arabidopsis thaliana
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1187)
  AUTHORS   Wu,K., Tian,L., Malik,K., Brown,D. and Miki,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-OCT-1999) Eastern Cereal and Oilseed Research Centre,
            Agriculture and Agri-Food Canada, Ottawa, Ontario K1A 0C6, Canada
FEATURES             Location/Qualifiers
     source          1. .1187
                     /organism="Arabidopsis thaliana"
                     /cultivar="Columbia"
                     /db_xref="taxon:3702"
                     /clone="133F15"
     gene            1. .1187
                     /gene="HD2B"
     CDS             61. .978
                     /gene="HD2B"
                     /codon_start=1
                     /product="putative histone deacetylase"
                     /protein_id="AAG28473.1"
                     /db_xref="GI:11066137"
                     /translation="MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLS
                     VTVGGAKLVIGTLSQDKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFT
                     SSDDEDVPEAVPAPAPTAVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDES
                     DDEDESEEDDDSEKGMDVDEDDSDDDEEEDSEDEEEEEFTPKKPEPINKKRPNESVSKT
                     PVSGKKAKPAAAPASTPQKTEKKKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGGN
                     NNKKPFNSGKQFGGSNNKGSNKGKGKRA"
BASE COUNT      377 a    234 c    288 g    288 t
ORIGIN


  Query Match           16.6%;  Score 156.2;  DB 8;  Length 1187;
  Best Local Similarity   64.1%;  Pred. No. 2.5e-24;
  Matches  252;  Conservative    0;  Mismatches  138;  Indels    3;  Gaps    1;

Qy      25 ttttctcaaccttgattcttagccatggagttctggggaattgaagttaaatcaggaaag 84
           || ||||  ||| | |    | | ||||||||||||||| ||| || | | ||  |||
Db      37 TTCTCTCTTCCTCGTTCAACAACAATGGAGTTCTGGGGAGTTGCGGTGACACCAAAAAAC 96

Qy      85 ccagttacagtgactcctgaagaaggcattcttatccacgtttctcaggcatcgcttgga 144
             |   ||  |||||||||||||||| ||  ||| ||||| |||||||||| || |||
Db      97 GCTACTAAGGTGACTCCTGAAGAAGACAGCCTTGTCCACATTTCTCAGGCTTCACTT--- 153

Qy     145 gaatgtaaaaacaagaaggqagaqtttgtgccttacatgtaaaggttgggaaccagaac 204
           || || | |   ||    |||| | |||  |||   ||| |  |||||        ||
Db     154 GACTGCACAGTGAAATCTGGAGAATCTGTGGTTTTGAGTGTGACTGTTGGTGGGGCTAAA 213

Qy     205 ttggttctgggaactctatcgactgagaacatccctcagcttttctgtgatttggtattc 264
           | ||| | ||||| || ||    || ||  ||||||| ||  || |||||||||  ||
Db     214 CTTGTTATTGGAACACTTTCACAAGACAAGTTCCCTCAGATTAGCTTTGATTTGGTTTTT 273

Qy     265 gacaaggagtttgagctttctcacacttggggaaaaggaagtgtttactttgttggatac 324
           || || ||||||||||||| ||||   |    |||| || |||| | ||  |||| |||
Db     274 GATAAAGAGTTTGAGCTTTCACACAGCGGTACCAAAGCAAATGTTCATTTCATTGGCTAC 333

Qy     325 aaaactcccaacattgagccacaaggctattctgaggaagaagaggaagaagaggaagaa 384
           ||| | |||||||| ||||   | | ||   ||       || || || || |     |||
Db     334 AAATCCCCCAACATCGAGCAGGATGACTTCACTAGTTCGGATGATGAGGATGTTCCTGAA 393

Qy     385 gttcctgctgggaatgctgccaaggctgtagct 417
           | |  | |||    |||  | |  |||||  ||
Db     394 GCTGTTCCTGCTCCTGCCCCTACTGCTGTTACT 426


RESULT   7
AX139020
LOCUS       AX139020               1218 bp    DNA     linear   PAT 30-MAY-2001
DEFINITION  Sequence 7 from Patent EP1094112.
ACCESSION   AX139020
VERSION     AX139020.1  GI:14274701
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1218)
  AUTHORS   Wu,K., Miki,B.L., Tian,L. and Brown,D.C.
  TITLE     Repressing gene expression in plants
  JOURNAL   Patent: EP 1094112-A 7 25-APR-2001;
            The Minister of Agriculture and Agri-Food (CA)
FEATURES             Location/Qualifiers
     source          1. .1218
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
BASE COUNT      395 a    237 c    296 g    290 t
ORIGIN


  Query Match           16.6%;  Score 156.2;  DB 6;  Length 1218;
  Best Local Similarity   64.1%;  Pred. No. 2.4e-24;
  Matches  252;  Conservative    0;  Mismatches  138;  Indels    3;  Gaps    1;

Qy      25 ttttctcaaccttgattcttagccatggagttctggggaattgaagttaaatcaggaaag 84
           || ||||  ||| | |    | | ||||||||||||||| ||| || | | ||  |||
Db      37 TTCTCTCTTCCTCGTTCAACAACAATGGAGTTCTGGGGAGTTGCGGTGACACCAAAAAAC 96

Qy      85 ccagttacagtgactcctgaagaaggcattcttatccacgtttctcaggcatcgcttgga 144
             |   ||  |||||||||||||||| ||  ||| ||||| |||||||||| || |||
Db      97 GCTACTAAGGTGACTCCTGAAGAAGACAGCCTTGTCCACATTTCTCAGGCTTCACTT--- 153

Qy     145 gaatgtaaaaacaagaaggqagaqtttgtgccttacatgtaaaggttgggaaccagaac 204
           || || | |   ||    |||| | |||  |||   ||| |  |||||        ||
Db     154 GACTGCACAGTGAAATCTGGAGAATCTGTGGTTTTGAGTGTGACTGTTGGTGGGGCTAAA 213

Qy     205 ttggttctgggaactctatcgactgagaacatccctcagcttttctgtgatttggtattc 264
           | ||| | ||||| || ||    || ||  ||||||| ||  || |||||||||  ||
Db     214 CTTGTTATTGGAACACTTTCACAAGACAAGTTCCCTCAGATTAGCTTTGATTTGGTTTTT 273

Qy     265 gacaaggagtttgagctttctcacacttggggaaaaggaagtgtttactttgttggatac 324
           || || ||||||||||||| ||||   |    |||| || |||| | ||  |||| |||
Db     274 GATAAAGAGTTTGAGCTTTCACACAGCGGTACCAAAGCAAATGTTCATTTCATTGGCTAC 333

Qy     325 aaaactcccaacattgagccacaaggctattctgaggaagaagaggaagaagaggaagaa 384
           ||| | |||||||| ||||   | | ||   ||       || || || || |     |||
Db     334 AAATCCCCCAACATCGAGCAGGATGACTTCACTAGTTCGGATGATGAGGATGTTCCTGAA 393

Qy     385 gttcctgctgggaatgctgccaaggctgtagct 417
           | |  | |||    |||  | |  |||||  ||
Db     394 GCTGTTCCTGCTCCTGCCCCTACTGCTGTTACT 426


RESULT   8
AF044914
```

```
Db  61414 TTGTTATAGCTTTCTGTGATTTGTCTCTTTTTGTTTACCATTTCAGAATAATGACTGAGG 61355

Qy    490 ------------------------------------------------------------ 489

Db  61354 AGATGAGACATCTTTTTTGAACAAATGTTGGTGACGCATTCTTGTGTTATGATTCTTGCA 61295

Qy    490 -atggatgaagatgattctgatggtgaggattctgaggaagaagagcctacacctaagaa 548
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61294 GATGGATGAAGATGATTCTGATGGTGAGGATTCTGAGGAAGAAGAGCCTACACCTAAGAA 61235

Qy    549 gcctgcatcaagcaagaagagagctaatgaaactacccctaaagcacctgtgtcagcaaa 608
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61234 GCCTGCATCAAGCAAGAAGAGAGCTAATGAAACTACCCCTAAAGCACCTGTGTCAGCAAA 61175

Qy    609 gaaggcgaaagtagcagttactcctcagaaa---------------------------- 639
           ||||||||||||||||||||||||||||
Db  61174 GAAGGCGAAAGTAGCAGTTACTCCTCAGAAAACAGGTATAGATACAAACCATTCTCTTAC 61115

Qy    640 ------------------------------------------------------------a 640
                                                                         |
Db  61114 ATTATTCACTAAATTGGGAAACACGAGTATTCATTTGAACCTGCATCGTCTTCTTTGGCA 61055

Qy    641 cagatgagaagaagaaaaggggggaaaggctgcaaaccagagcccaaagtcggccagtcaag 700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61054 CAGATGAGAAGAAGAAAGGGGGGAAAGGCTGCAAACCAGAGCCCAAAGTCGGCCAGTCAAG 60995

Qy    701 tctcatgtggttcatgcaagaa---------------------------------- 722
           ||||||||||||||||||||||
Db  60994 TCTCATGTGGTTCATGCAAGAAGTGAGACTCTCTCTCTTTCTCTAACTCTTGTCTTCACT 60935

Qy    723 ------------------------------------------------------gac 725
                                                                   |||
Db  60934 GCTCAAAATTTCTGTTTGTTATAACATTTCGCTTGTTGGCTTATATCTCTTTGGCAGGAC 60875

Qy    726 tttcaactcagggaatgcacttgagtctcacaacaaggccaagcacgctgctgccaagtg 785
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  60874 TTTCAACTCAGGGAATGCACTTGAGTCTCACAACAAGGCCAAGCACGCTGCTGCCAAGTG 60815

Qy    786 aagtggtttcttattagagcttgtgatttctatggaattttgcctgtagtctttatgaaa 845
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  60814 AAGTGGTTTCTTATTAGAGCTTGTGATTTCTATGGAATTTTGCCTGTAGTCTTTATGAAA 60755

Qy    846 ccttcggatttttcttatatttttcttttgataacaagagtcttaatgaaagagagccagtt 905
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  60754 CCTTCGGATTTTCTTATATTTTCTTTTGATAACAAGAGTCTTAATGAAAGAGAGCCAGTT 60695

Qy    906 ggagtctta 914
           |||||||||
Db  60694 GGAGTCTTA 60686


RESULT  5
AY059893
LOCUS       AY059893               1172 bp    mRNA    linear   PLN 04-NOV-2001
DEFINITION  Arabidopsis thaliana histone deacetylase-like protein (At5g22650;
            MDJ22.7) mRNA, complete cds.
ACCESSION   AY059893
VERSION     AY059893.1  GI:16649046
KEYWORDS    FLI_CDNA.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1172)
  AUTHORS   Nguyen,M., Karlin-Neumann,G., Southwick,A., Lam,B., Miranda,M.,
            Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,
            Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
            Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
            Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,
            Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-OCT-2001) DNA Sequencing and Technology Center,
            Stanford University, 855 California Avenue, Palo Alto, CA 94304,
            USA
COMMENT     e-mail for correspondence: arab@sequence.stanford.edu

            RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis  Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members carried out the
            sequencing and annotation of the RAFL cDNAs: Nguyen,M.,
            Southwick,A., Karlin-Neumann,G.,  Lam,B., Miranda,M., Palm,C.J.,
            Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K.,
            Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P.,
            Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

            Nguyen,M, (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed
            equally to this work.  Shinozaki,K. (RIKEN GSC) and Davis,R.W.
            (SSP/Stanford)  contributed equally to this work as PIs.
FEATURES             Location/Qualifiers
     source          1. .1172
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /chromosome="V"
                     /clone="RAFL08-18-E14"
                     /note="This clone is in pBluescript"
     gene            1. .1172
                     /gene="At5g22650; MDJ22.7"
     CDS             75. .995
                     /gene="At5g22650; MDJ22.7"
                     /codon_start=1
                     /product="histone deacetylase-like protein"
                     /protein_id="AAL24375.1"
                     /db_xref="GI:16649047"
                     /translation="MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLS
                     VTVGGAKLVIGTLSQDKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFT
                     SSDDEDVPEAVPAPAPTAVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDES
                     DDEDFSEEDDDSEKGMDVDEDDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKT
                     PVSGKKAKPAAAPASTPQKTEEKKKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGG
                     NNNKKPFNSGKQFGGSNNKGSNKGKGKGRA"
BASE COUNT      362 a    234 c    285 g    291 t
ORIGIN


  Query Match             16.6%;  Score 156.2;  DB 8;  Length 1172;
  Best Local Similarity   64.1%;  Pred. No. 2.5e-24;
  Matches  252;  Conservative    0;  Mismatches  138;  Indels    3;  Gaps    1;

Qy      25 ttttctcaaccttgattcttagccatggagttctggggaattgaagttaaatcaggaaag 84
           || ||||  ||| | |    | | ||||||||||||||| |||  || | | ||  |||
Db      51 TTCTCTCTTCCTCGTTCAACAACAATGGAGTTCTGGGGAGTTGCGGTGACACCAAAAAAC 110

Qy      85 ccagttacagtgactcctgaagaaggcattcttatccacgtttctcaggcatcgcttgga 144
             |   ||  |||||||||||||||| ||  ||| ||||| |||||||||| || |||
Db     111 GCTACTAAGGTGACTCCTGAAGAAGACAGCCTTGTCCACATTTCTCAGGCTTCACTT--- 167

Qy     145 gaatgtaaaaacaagaagggagagtttgtgccttacatgtaaaggttgggaaccagaac 204
           || || | |   ||    ||||| | ||||  |||   ||| |  |||||       ||
Db     168 GACTGCACAGTGAAATCTGGAGAATCTGTGGTTTTGAGTGTGACTGTTGGTGGGGCTAAA 227

Qy     205 ttggttctgggaactctatcgactgagaacatccctcagcttttctgtgatttggtattc 264
             | ||| | ||||| || ||    || ||  |||||||| ||  || |||||||| ||
Db     228 CTTGTTATTGGAACACTTTCACAAGACAAGTTCCCTCAGATTAGCTTTGATTTGGTTTTT 287

Qy     265 gacaaggaqtttgaqctttctcacacttggggaaaaggaaqtgtttactttgttggatac 324
           || || ||||||||||||||| ||||   |    |||| || |||| | ||  |||| |||
Db     288 GATAAAGAGTTTGAGCTTTCACACAGCGGTACCAAAGCAAATGTTCATTTCATTGGCTAC 347

Qy     325 aaaactcccaacattgagccacaaggctattctgaggaagaagaggaagaagaggaagaa 384
           ||| | ||||||||| ||||   | | ||   ||       || || || || |    |||
Db     348 AAATCCCCCAACATCGAGCAGGATGACTTCACTAGTTCGGATGATGAGGATGTTCCTGAA 407
```

```
                similar to chromosome V (7873-8427), indirect repeat (copy
                B, 8428-8440), direct repeat (copy B, 8441-8443)"
                /note="Region: Putative transposon"
                /evidence=not_experimental
repeat_region   7860. .7872
                /note="inverted repeat flanking putative transposon from
                position 7873 to 8427, copy A"
                /rpt_type=inverted
misc_feature    7860. .8445
                /note="These 586 nucleotides are 97.9% identical to a
                region of chromosome V clone MBK5 (AB005234), position
                81694 to 82277 and encompass the flanking inverted
                repeats"
                /note="Region: Similar to clone MBK5"
repeat_region   complement(8428. .8440)
                /note="inverted repeat flanking putative transposon from
                position 7873 to 8427, copy B"
                /rpt_type=inverted
repeat_region   8441. .8443
                /note="direct repeat flanking putative transposon from
                position 7873 to 8427, copy B"
                /rpt_type=direct
repeat_region   8575. .8700
                /note="inverted repeat 1, copy A; copy B located from 9649
                to 9775"
                /rpt_type=inverted
repeat_region   8872. .9139
                /note="direct repeat 1, copy A; copy B located from 9789
                to 10041"
                /rpt_type=direct
exon            complement(9013. .9156)
                /note="predicted by Grail"
                /evidence=not_experimental
exon            complement(9446. .9741)
                /note="predicted by Grail"
                /evidence=not_experimental
exon            complement(9477. .9741)
                /note="predicted by MZEF; shows moderately weak similarity
                to EST F14401"
                /evidence=not_experimental
repeat_region   complement(9649. .9775)
                /note="inverted repeat 1, copy B; copy A located from 8575
                to 8700"
                /rpt_type=inverted
repeat_region   9789. .10041
                /note="direct repeat 1, copy B; copy A located from 8872
                to 9139"
                /rpt_type=direct
exon            complement(9819. .10293)
                /note="predicted by MZEF; shows moderately weak similarity
                to EST F14401"
                /evidence=not_experimental
exon            complement(9819. .10054)
                /note="predicted by Grail"
                /evidence=not_experimental
misc_feature    12951. .13621
                /note="region shares 90.3% identity to chromosome V BAC
                T21B04 (AF007271) from position 32026 to 32690 on the
                reverse strand"
                /note="Region: Similarity to BAC T21B04"
gene            complement(14523. .17066)
                /gene="T32N15.3"
                /note="gene encoding hypothetical protein similar to
                T32N15.4"
                /evidence=not_experimental
CDS             complement(join(14523. .14849,15006. .15264,15584. .16038,
                16325. .16643,16706. .16771,16969. .17066))
                /gene="T32N15.3"
                /note="T32N15.3 is most similar to hypothetical protein
                T32N15.4 (most nucleotide sequence differences are second
                or third position changes separated by 3 to 5 bases of
                identity, a hallmark of a multigene family rather than a
                transposon). T32N15.3 shows moderate similarity to the
                conceptual translations of A. thaliana ESTs F14401 and
                F14402. These ESTs are the 5' and 3' sequences,
                respectively, of clone TAP0138."
                /codon_start=1
                /evidence=not_experimental
                /product="hypothetical protein"
                /protein_id="AAB70027.1"
                /db_xref="GI:2392764"
                /translation="MGYEKFTWVIKNFSSIQSEYIKSDIFVIGGCKWRIRFRLTVVNQ
                LSENLSRRGGDFVFLWSTEGKRWFDKKLPLCGYEEVLLLTKLNAKHGGFLVNNEVK1V
                AEVDVLEVIGKLDVSKESQEVIKPRKRMRLYGDGGAVSSHLHKETSSVDVNGFQVLPS
                QAESVKRIFERHPYMALEFRAKNQQLRASCINVLLSLIKTLCQSLQDISIDDLGQTEQ
                VLTFLQNSGFKVDWLERKLEEVKEKKIQEHIGKSRMQGLEEDLKVFKKKCSDIEALLE
                KEKEELKGLKQKCSDIEALLEKEKGKVLAAAARTPLTLDRMNILIGYREMISLTKLNA
                KKGGFLVNNKVKIVVEVDVLQVVGGLDVSQEVTQPLKRIRLNDDGAASTHLQTIDVNG
                FLVLASQVESVKRIFESHPEMTLTFRANNQQLRTSCINVLLNLIETMCMSPQNLSIAD
                LGQTEQALTYLKNSGFKVDWLERKLEEVTEKKIQEHIGKSRMQGLEEDLKVFKKNLL"
repeat_region   complement(18372. .19531)
                /note="direct repeat 2, copy A"
                /rpt_type=direct
misc_feature    complement(18372. .26641)
                /note="putative del-like retrotransposon; sequence from
                20428 to 24996 is similar to the sequence encoding the del
                multifunctional protein of Henry's lily (X13886); this
                sequence is flanked by long direct repeats of 1.16 kbp"
                /note="Region: Putative del-like retrotransposon"
                /evidence=not_experimental
repeat_region   complement(25476. .26641)
                /note="direct repeat 2, copy B"
                /rpt_type=direct
misc_feature    complement(26833. .29405)
```

Query Match 26.5%; Score 249; DB 8; Length 101371;
Best Local Similarity 61.8%; Pred. No. 4.6e-45;
Matches 784; Conservative 0; Mismatches 0; Indels 485; Gaps 5;

```
Qy    131 aggcatcgcttggagaatgtaaaaacaagaagggagagtttgtgccttacatgtaaagg 190
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61954 AGGCATCGCTTGGAGAATGTAAAAACAAGAAGGGAGAGTTTGTGCCTTTACATGTAAAGG 61895

Qy    191 ttgggaaccagaacttggttctgggaactctatcgactgagaacatccctcagcttttct 250
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61894 TTGGGAACCAGAACTTGGTTCTGGGAACTCTATCGACTGAGAACATCCCTCAGCTTTTCT 61835

Qy    251 gtgatttggtattcgacaaggagtttgagctttctcacacttggggaaaaggaagtgttt 310
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61834 GTGATTTGGTATTCGACAAGGAGTTTGAGCTTTCTCACACTTGGGGAAAAGGAAGTGTTT 61775

Qy    311 actttgttggatacaaaactcccaacattgagccacaaggctattct---------- 357
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61774 ACTTTGTTGCATACAAAACTCCCAACATTGAGCCACAAGGCTATTCTGGTGAGTTTGCAA 61715

Qy    358 ---------------------------------------------------------- 357

Db  61714 GATAAAATCTGTTTCAATTCAATTATTCTGTTTGTAACTGTAATATTGGCTTGGTTGCTG 61655

Qy    358 --------------------------------gaggaagaagaggaagaagaggaagaa 384
                                           ||||||||||||||||||||||||||||
Db  61654 ATGATTCGTTTATGTGTTGTATTGATATATTCAGAGGAAGAAGAGGAAGAAGAGGAAGAA 61595

Qy    385 gttcctgct---------------------------------------------------- 393
          |||||||||
Db  61594 GTTCCTGCTGGTACTTAAACTAGACAGAGACTACTTGATTTGTGATCCTGATAATTGGTA 61535

Qy    394 -------------------gggaatgctgccaaggctgtagctaaaccaaaggctaagcc 434
                              ||||||||||||||||||||||||||||||||||||||||||
Db  61534 ATAACTAGTTCCATTTGCAGGGAATGCTGCCAAGGCTGTAGCTAAACCAAAGGCTAAGCC 61475

Qy    435 tgcagaagtgaagccagctgttgatgatgaagaggatgagtctgattctgacgga----- 489
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61474 TGCAGAAGTGAAGCCAGCTGTTGATGATGAAGAGGATGAGTCTGATTCTGACGGAGTGAG 61415

Qy    490 ---------------------------------------------------------- 489
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    262 TCTGTGATTTGGTATTCGACAAGGAGTTTGAGCTTTCTCACACTTGGGGAAAAGGAAGTG 321

Qy    308 tttactttgttggatacaaaactcccaacattgagccacaaggctattctgaggaagaag 367
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    322 TTTACTTTGTTGGATACAAAACTCCCAACATTGAGCCACAAGGCTATTCTGAGGAAGAAG 381

Qy    368 aggaagaagaggaagaagttcctgctgggaatgctgccaaggctgtagctaaaccaaagg 427
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    382 AGGAAGAAGAGGAAGAAGTTCCTGCTGGGAATGCTGCCAAGGCTGTAGCTAAACCAAAGG 441

Qy    428 ctaagcctgcagaagtgaagccagctgttgatgatgaagaggatgagtctgattctgacg 487
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    442 CTAAGCCTGCAGAAGTGAAGCCAGCTGTTGATGATGAAGAGGATGAGTCTGATTCTGACG 501

Qy    488 gaatggatgaagatgattctgatggtgaggattctgaggaagaagagcctacacctaaga 547
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    502 GAATGGATGAAGATGATTCTGATGGTGAGGATTCTGAGGAAGAAGAGCCTACACCTAAGA 561

Qy    548 agcctgcatcaagcaagaagagagctaatgaaactacccctaaagcacctgtgtcagcaa 607
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    562 AGCCTGCATCAAGCAAGAAGAGAGCTAATGAAACTACCCCTAAAGCACCTGTGTCAGCAA 621

Qy    608 agaaggcgaaagtagcagttactcctcagaaaacagatgagaagaagaaaggggggaaagg 667
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    622 AGAAGGCGAAAGTAGCAGTTACTCCTCAGAAAACAGATGAGAAGAAGAAAGGGGGAAAGG 681

Qy    668 ctgcaaaccagagcccaaagtcggccagtcaagtctcatgtggttcatgcaagaagactt 727
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    682 CTGCAAACCAGAGCCCAAAGTCGGCCAGTCAAGTCTCATGTGGTTCATGCAAGAAGACTT 741

Qy    728 tcaactcagggaatgcacttgagtctcacaacaaggccaagcacgctgctgccaagtgaa 787
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    742 TCAACTCAGGGAATGCACTTGAGTCTCACAACAAGGCCAAGCACGCTGCTGCCAAGTGAA 801

Qy    788 gtggtttcttattagagcttgtgatttctatggaattttgcctgtagtctttatgaaacc 847
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    802 GTGGTTTCTTATTAGAGCTTGTGATTTCTATGGAATTTTGCCTGTAGTCTTTATGAAACC 861

Qy    848 ttcggatttttcttatattttctttttgataacaagagtcttaatgaaagagagccagttgg 907
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    862 TTCGGATTTTCTTATATTTTCTTTTGATAACAAGAGTCTTAATGAAAGAGAGCCAGTTGG 921

Qy    908 agtctta 914
          |||||||
Db    922 AGTCTTA 928


RESULT  4
AC002534/c
LOCUS       AC002534              101371 bp    DNA     linear   PLN 21-APR-1999
DEFINITION  Arabidopsis thaliana BAC T32N15 from chromosome III near 54 cM,
            complete sequence.
ACCESSION   AC002534
VERSION     AC002534.1  GI:2392762
KEYWORDS    HTG.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 101371)
  AUTHORS   de la Bastide,M.R., Parnell,L.D., Kaplan,N., Gnoj,L., Hameed,A.,
            Schutz,K., Hasegawa,A., Gottesman,T., Shohdy,N., Granat,S.,
            Jensen,K., Johnson,A.F., Lodhi,M., Dedhia,N., Martienssen,R. and
            McCombie,W.R.
  TITLE     A. thaliana BAC T32N15 from chromosome V
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 101371)
  AUTHORS   McCombie,W.R. and de la Bastide,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-SEP-1997) Lita Annenberg Hazen Genome Sequencing
            Center, Cold Spring Harbor Laboratory, 1 Bungtown Rd., Cold Spring
            Harbor, NY 11724, USA
REFERENCE   3  (bases 1 to 101371)
  AUTHORS   Parnell,L.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-APR-1999) Lita Annenberg Hazen Genome Sequencing
            Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring
            Harbor, NY 11724
  REMARK    Arabidopsis thaliana BAC T32N15 from chromosome III near 54 cM
COMMENT     Sequencing, analysis, and annotation were performed in
            collaboration with the CSHL/WU/ABI Arabidopsis Sequencing
            Consortium. Information on physical mapping, YAC and BAC library
            construction, and a more detailed annotation of the seqence can be
            viewed at http://www.cshl.org/arabweb/. Gene prediction programs
            Grail and MZEF were used in the assembly of coding regions. In
            certain regions of the BAC these exons are not incorporated into a
            good model gene and are thus annotated as single exons. BAC T32N15
            has been mapped to chromosome III near 54 cM and marker m249. A
            graphical display of the annotation of T32N15 can be found at
            http://www.cshl.org/arabweb/T32N15-titlepage.html.
FEATURES             Location/Qualifiers
     source          1. .101371
                     /organism="Arabidopsis thaliana"
                     /cultivar="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="V"
                     /map="near 54 cM"
     gene            <404. .1903
                     /gene="T32N15.1"
                     /note="gene encoding hypothetical protein; 5' region of
                     T32N15.1 is not likely to be encoded by BAC T32N15"
                     /evidence=not_experimental
     CDS             join(<404. .1327,1413. .1559,1616. .1903)
                     /gene="T32N15.1"
                     /note="amino-terminus is not likely to be encoded by BAC
                     T32N15; sequence is similar to many entries of A. thaliana
                     genomic sequence, for example MX110, GenBank accession
                     number AB005248"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="hypothetical protein"
                     /protein_id="AAB70026.1"
                     /db_xref="GI:2392763"
                     /translation="VPLGMSSVSTASLNCLPDELLVHVLSSLETKQAASTSVLSKRWR
                     TLFAVRRNLDFDDSIISHPEVGEQNMDDVQESFRDFVDKRLAFQGSVPINKFSLIYGD
                     KHDDVRVDRWINTALEHGVSELHLCLTSVTRRLHRFPSNVFRSTTLVKLTLGTNLFIV
                     YFPSDTCLPVLKILVLDSIWFDRIKFSNVLIAGCPALEDLTIDQKSFPGLPNVVSSKT
                     VKSLSIVYKYSADFDWFRTVALDTPNLVTLLYSTYARHRYRHCNLESLVNATLDLHFL
                     ENCDEAFEPNVTDLMIAVRNVQMLHLTSSATEVISQCCKGGLPMFKNLLVLVFLGNTE
                     RVWKVFLPLLLEHSPNLTKLCLESLFLILQGLYHGTDEDEFDEIHIPRSNKVNMLRII
                     QCQGTENELKHISHFLLKMECLQLVQVNFSETIVDSKKVQLTEDLMKLPSASSRLTMQ
                     VI"
     repeat_region   3559. .3750
                     /note="function=putative_MITES; region is 96.9% identical
                     to the nucleotide sequence of A. thaliana BAC T7N9,
                     GenBank accession number AC000348; although GRAIL predicts
                     an exon from position 3620 to 3759 and translation of this
                     exon is nearly identical to protein T7N9.14;
                     gene-prediction programs fail to assemble this entire exon
                     into an mRNA sequence"
                     /rpt_type=dispersed
     repeat_region   6329. .6981
                     /note="function=minisatellite; 23 copies of a 28-bp tandem
                     repeat"
                     /rpt_type=dispersed
                     /rpt_unit=TTTTCCCGCCAAAAACNTAAACCCGTRA
     repeat_region   7857. .7859
                     /note="function=direct_repeat; cognate A of direct repeat
                     flanking putative transposon from position 7873 to 8427"
                     /rpt_type=direct
     misc_feature    7857. .8443
                     /note="putative non-autonomous transposon; genomic
                     organization is: direct repeat (copy A, position
                     7857-7859), indirect repeat (copy A, 7860-7872), region
```

```
Db    361 GAAGAAGAGGAAGAAGAGGAAGAAGTTCCTGCTGGGAATGCTGCCAAGGCTGTAGCTAAA 420

Qy    421 ccaaaggctaagcctgcagaagtgaagccagctgttgatgatgaagaggatgagtctgat 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CCAAAGGCTAAGCCTGCAGAAGTGAAGCCAGCTGTTGATGATGAAGAGGATGAGTCTGAT 480

Qy    481 tctgacggaatggatgaagatgattctgatggtgaggattctgaggaagaagagcctaca 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TCTGACGGAATGGATGAAGATGATTCTGATGGTGAGGATTCTGAGGAAGAAGAGCCTACA 540

Qy    541 cctaagaagcctgcatcaagcaagaagagagctaatgaaactacccctaaagcacctgtg 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CCTAAGAAGCCTGCATCAAGCAAGAAGAGAGCTAATGAAACTACCCCTAAAGCACCTGTG 600

Qy    601 tcagcaaagaaggcgaaagtagcagttactcctcagaaaacagatgagaagaagaaaggg 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TCAGCAAAGAAGGCGAAAGTAGCAGTTACTCCTCAGAAAACAGATGAGAAGAAGAAAGGG 660

Qy    661 ggaaaggctgcaaaccagagcccaaagtcggccagtcaagtctcatgtggttcatgcaag 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GGAAAGGCTGCAAACCAGAGCCCAAAGTCGGCCAGTCAAGTCTCATGTGGTTCATGCAAG 720

Qy    721 aagactttcaactcagggaatgcacttgagtctcacaacaaggccaagcacgctgctgcc 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 AAGACTTTCAACTCAGGGAATGCACTTGAGTCTCACAACAAGGCCAAGCACGCTGCTGCC 780

Qy    781 aagtgaagtggtttcttattagagcttgtgatttctatggaattttgcctgtagtcttta 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AAGTGAAGTGGTTTCTTATTAGAGCTTGTGATTTCTATGGAATTTTGCCTGTAGTCTTTA 840

Qy    841 tgaaaccttcggatttttcttatattttctttttgataacaagagtcttaatgaaagagagc 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TGAAACCTTCGGATTTTCTTATATTTTCTTTTGATAACAAGAGTCTTAATGAAAGAGAGC 900

Qy    901 caqttggagtctta 914
          ||||||||||||||
Db    901 CAGTTGGAGTCTTA 914
```

```
RESULT   3
AY065396
LOCUS       AY065396                977 bp    mRNA    linear   PLN 11-DEC-2001
DEFINITION  Arabidopsis thaliana putative histone deacetylase (At3g44750) mRNA,
            complete cds.
ACCESSION   AY065396
VERSION     AY065396.1  GI:17529219
KEYWORDS    FLI_CDNA.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 977)
  AUTHORS   Yamada,K., Banh,J., Chang,C.H., Chang,E., Dale,J.M.,
            Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.,
            Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Yu,S., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
            Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
            Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Full Length cDNA of gene At3g44750 (GI:15230496)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 977)
  AUTHORS   Yamada,K., Banh,J., Chang,C.H., Chang,E., Dale,J.M.,
            Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C.,
            Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Yu,S., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
            Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
            Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-DEC-2001) Plant Gene Expression Center, 800 Buchanan
            Street, Albany, CA 94710, USA
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members carried out the
            sequencing and annotation of the RAFL cDNAs: Yamada,K., Banh,J.,
            Chang,C.H., Chang,E., Dale,J.M., Goldsmith,A.D., Lee,J.M.,
            Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C.,
            Yamamura,Y., Yu,G., Yu,S., Bowser,L., Chen,H., Cheuk,R., Jones,T.,
            Karlin-Neumann,G., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M.,
            Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Davis,R.W.,
            Ecker,J.R. and Theologis,A.

            Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to
            this work.  Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC)
            contributed equally to this work as PIs.
FEATURES             Location/Qualifiers
     source          1. .977
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /chromosome="3"
                     /clone="RAFL06-70-K17 (R11479)"
                     /note="This clone is in a modified pBluescript vector
                     (Lambda ZAP) as a XhoI/SstI insert.
                     ecotype: Columbia"
     gene            1. .977
                     /gene="At3g44750"
     5'UTR           1. .62
                     /gene="At3g44750"
     CDS             63. .800
                     /gene="At3g44750"
                     /codon_start=1
                     /evidence=experimental
                     /product="putative histone deacetylase"
                     /protein_id="AAL38837.1"
                     /db_xref="GI:17529220"
                     /translation="MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPL
                     HVKVGNQNLVLGTLSTENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGY
                     SEEEEEEEEEVPAGNAAKAVAKPKAKPAEVKPAVDDEEDESDSDGMDEDDSDGEDSEE
                     EEPTPKKPASSKKRANETTPKAPVSAKKAKVAVTPQKTDEKKKGGKAANQSPKSASQV
                     SCGSCKKTPNSGNALESHNKAKHAAAK"
     3'UTR           801. .977
                     /gene="At3g44750"
BASE COUNT      311 a    183 c    231 g    252 t
ORIGIN
```

```
  Query Match             96.4%;  Score 905.4;  DB 8;  Length 977;
  Best Local Similarity   99.9%;  Pred. No. 1.3e-189;
  Matches  906;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      8 ccgtaaaaatcctctcttttttctcaaccttgattcttagccatggagttctggggaattg 67
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     22 CCCTAAAAATCCTCTCTTTTTCTCAACCTTGATTCTTAGCCATGGAGTTCTGGGGAATTG 81

Qy     68 aagttaaatcaggaaagccagttacagtgactcctgaagaaggcattcttatccacgttt 127
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     82 AAGTTAAATCAGGAAAGCCAGTTACAGTGACTCCTGAAGAAGGCATTCTTATCCACGTTT 141

Qy    128 ctcaggcatcgcttggagaatgtaaaaacaagaagggagagtttgtgcctttacatgtaa 187
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    142 CTCAGGCATCGCTTGGAGAATGTAAAAACAAGAAGGGAGAGTTTGTGCCTTTACATGTAA 201

Qy    188 aggttgggaaccagaacttggttctgggaactctatcgactgagaacatccctcagcttt 247
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    202 AGGTTGGGAACCAGAACTTGGTTCTGGGAACTCTATCGACTGAGAACATCCCTCAGCTTT 261

Qy    248 tctgtgatttggtattcgacaaggagtttgaggctttctcacacttgggggaaaaggaaqtg 307
```

```
  Query Match             100.0%;  Score 939;  DB 6;  Length 939;
  Best Local Similarity   100.0%;  Pred. No. 5.2e-197;
  Matches  939;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 cacgcgtccgtaaaaatcctctctttttctcaaccttgattcttagccatggagttctgg 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CACGCGTCCGTAAAAATCCTCTCTTTTTCTCAACCTTGATTCTTAGCCATGGAGTTCTGG 60

Qy      61 ggaattgaagttaaatcaggaaagccagttacagtgactcctgaagaaggcattcttatc 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GGAATTGAAGTTAAATCAGGAAAGCCAGTTACAGTGACTCCTGAAGAAGGCATTCTTATC 120

Qy     121 cacgtttctcaggcatcgcttggagaatgtaaaaacaagaagggagagtttgtgccttta 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CACGTTTCTCAGGCATCGCTTGGAGAATGTAAAAACAAGAAGGGAGAGTTTGTGCCTTTA 180

Qy     181 catgtaaaggttgggaaccagaacttggttctgggaactctatcgactgagaacatccct 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CATGTAAAGGTTGGGAACCAGAACTTGGTTCTGGGAACTCTATCGACTGAGAACATCCCT 240

Qy     241 cagcttttctgtgatttggtattcgacaaggagtttgagctttctcacacttggggaaaa 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CAGCTTTTCTGTGATTTGGTATTCGACAAGGAGTTTGAGCTTTCTCACACTTGGGGAAAA 300

Qy     301 ggaagtgtttacttttgttggatacaaaaactcccaacattgaagccacaaggctattctgag 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GGAAGTGTTTACTTTGTTGGATACAAAACTCCCAACATTGAGCCACAAGGCTATTCTGAG 360

Qy     361 gaagaagaggaagaagaggaagaagttcctgctgggaatgctgccaaggctgtagctaaa 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GAAGAAGAGGAAGAAGAGGAAGAAGTTCCTGCTGGGAATGCTGCCAAGGCTGTAGCTAAA 420

Qy     421 ccaaaggctaagcctgcagaagtgaagccagctgttgatgatgaagaggatgagtctgat 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CCAAAGGCTAAGCCTGCAGAAGTGAAGCCAGCTGTTGATGATGAAGAGGATGAGTCTGAT 480

Qy     481 tctgacggaatggatgaagatgattctgatggtgaggattctgaggaagaagaagcctaca 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TCTGACGGAATGGATGAAGATGATTCTGATGGTGAGGATTCTGAGGAAGAAGAGCCTACA 540

Qy     541 cctaagaagcctgcatcaagcaagaagagagctaatgaaactacccctaaagcacctgtg 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 CCTAAGAAGCCTGCATCAAGCAAGAAGAGAGCTAATGAAACTACCCCTAAAGCACCTGTG 600

Qy     601 tcagcaaagaaggcgaaagtagcagttactcctcagaaaacagatgagaagaagaaaaggg 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 TCAGCAAAGAAGGCGAAAGTAGCAGTTACTCCTCAGAAAACAGATGAGAAGAAGAAAGGG 660

Qy     661 ggaaaggctgcaaaccagagcccaaagtcggccagtcaagtctcatgtggttcatgcaag 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GGAAAGGCTGCAAACCAGAGCCCAAAGTCGGCCAGTCAAGTCTCATGTGGTTCATGCAAG 720

Qy     721 aagactttcaactcagggaatgcacttgagtctcacaacaaggccaagcacgctgctgcc 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 AAGACTTTCAACTCAGGGAATGCACTTGAGTCTCACAACAAGGCCAAGCACGCTGCTGCC 780

Qy     781 aagtgaagtggtttcttattagagcttgtgatttctatggaattttgcctgtagtctttta 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 AAGTGAAGTGGTTTCTTATTAGAGCTTGTGATTTCTATGGAATTTTGCCTGTAGTCTTTA 840

Qy     841 tgaaaccttcggatttttcttatattttcttttgataacaagagtcttaatgaaagagagc 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 TGAAACCTTCGGATTTTCTTATATTTTCTTTTGATAACAAGAGTCTTAATGAAAGAGAGC 900

Qy     901 cagttggagtcttaaaaaaaaaaaaaaaaaggggcggccgc 939
           |||||||||||||||||||||||||||||||||||||||||
Db     901 CAGTTGGAGTCTTAAAAAAAAAAAAAAAAAGGGCGGCCGC 939


RESULT   2
AF195545
```

```
LOCUS       AF195545                914 bp    mRNA    linear   PLN 01-NOV-2000
DEFINITION  Arabidopsis thaliana putative histone deacetylase (HD2A) mRNA,
            complete cds.
ACCESSION   AF195545
VERSION     AF195545.1  GI:11066134
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 914)
  AUTHORS   Wu,K., Tian,L., Malik,K., Brown,D. and Miki,B.
  TITLE     Functional analysis of HD2 histone deacetylase homologs in
            Arabidopsis thaliana
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 914)
  AUTHORS   Wu,K., Tian,L., Malik,K., Brown,D. and Miki,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-OCT-1999) Eastern Cereal and Oilseed Research Centre,
            Agriculture and Agri-Food Canada, Ottawa, Ontario K1A 0C6, Canada
FEATURES             Location/Qualifiers
     source          1. .914
                     /organism="Arabidopsis thaliana"
                     /cultivar="Columbia"
                     /db_xref="taxon:3702"
                     /clone="211N2"
     gene            1. .914
                     /gene="HD2A"
     CDS             49. .786
                     /gene="HD2A"
                     /codon_start=1
                     /product="putative histone deacetylase"
                     /protein_id="AAG28472.1"
                     /db_xref="GI:11066135"
                     /translation="MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPL
                     HVKVGNQNLVLGTLSTENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGY
                     SEEEEEEEEEVPAGNAAKAVAKPKAKPAEVKPAVDDEEDESDSDGMDEDDSDGEDSEE
                     EEPTPKKPASSKKRANETTPKAPVSAKKAKVAVTPQKTDEKKKGGKAANQSPKSASQV
                     SCGSCKKTFNSGNALESHNKAKHAAAK"
BASE COUNT      281 a    171 c    230 g    232 t
ORIGIN
```

```
  Query Match             97.3%;  Score 914;  DB 8;  Length 914;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-191;
  Matches  914;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 cacgcgtccgtaaaaatcctctctttttctcaaccttgattcttagccatggagttctgg 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CACGCGTCCGTAAAAATCCTCTCTTTTTCTCAACCTTGATTCTTAGCCATGGAGTTCTGG 60

Qy      61 ggaattgaagttaaatcaggaaagccagttacagtgactcctgaagaaggcattcttatc 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GGAATTGAAGTTAAATCAGGAAAGCCAGTTACAGTGACTCCTGAAGAAGGCATTCTTATC 120

Qy     121 cacgtttctcaggcatcgcttggagaatgtaaaaacaagaagggagagtttgtgccttta 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CACGTTTCTCAGGCATCGCTTGGAGAATGTAAAAACAAGAAGGGAGAGTTTGTGCCTTTA 180

Qy     181 catgtaaaggttgggaaccagaacttggttctgggaactctatcgactgagaacatccct 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CATGTAAAGGTTGGGAACCAGAACTTGGTTCTGGGAACTCTATCGACTGAGAACATCCCT 240

Qy     241 cagcttttctgtgatttggtattcgacaaggagtttgagctttctcacacttggggaaaa 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CAGCTTTTCTGTGATTTGGTATTCGACAAGGAGTTTGAGCTTTCTCACACTTGGGGAAAA 300

Qy     301 ggaagtgtttacttttgttggatacaaaaactcccaacattgaagccacaaggctattctgag 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GGAAGTGTTTACTTTGTTGGATACAAAACTCCCAACATTGAGCCACAAGGCTATTCTGAG 360

Qy     361 gaagaagaggaagaagaggaagaagttcctgctgggaatgctgccaaggctgtagctaaa 420
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 15:56:26 ; Search time 5336.92 Seconds
(without alignments)
3681.902 Million cell updates/sec

Title: US-09-645-337-5
Perfect score: 939
Sequence: 1 cacgcgtccgtaaaaatcct.....aaaaaaaaaagggcggccgc 939

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

MEHTA
09/645,337
SEQ ID: 5-8

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	939	100.0	939	6	AX139018	AX139018 Sequence
2	914	97.3	914	8	AF195545	AF195545 Arabidops
3	905.4	96.4	977	8	AY065396	AY065396 Arabidops
c 4	249	26.5	101371	8	AC002534	AC002534 Arabidops
5	156.2	16.6	1172	8	AY059893	AY059893 Arabidops
6	156.2	16.6	1187	8	AF195546	AF195546 Arabidops
7	156.2	16.6	1218	6	AX139020	AX139020 Sequence
8	149.8	16.0	918	8	AF044914	AF044914 Arabidops
9	124.2	13.2	1236	8	AF372889	AF372889 Arabidops
10	123.8	13.2	1307	6	AR168371	AR168371 Sequence
11	123.4	13.1	885	8	AF255712	AF255712 Arabidops
12	117.6	12.5	894	8	AF255711	AF255711 Oryza sat
13	109.2	11.6	1124	8	AF384033	AF384033 Zea mays
14	104.2	11.1	1245	6	AR168370	AR168370 Sequence
15	103.8	11.1	912	8	AF254072	AF254072 Zea mays
16	100	10.6	1191	6	AR168369	AR168369 Sequence
17	98.4	10.5	903	8	AF254073	AF254073 Zea mays
18	98.4	10.5	1283	6	AR168368	AR168368 Sequence
19	96.8	10.3	1169	8	ZMU82815	U82815 Zea mays nu
20	95	10.1	77363	8	AB006699	AB006699 Arabidops
c 21	85.6	9.1	132323	8	AP003416	AP003416 Oryza sat
22	85.6	9.1	174218	2	AC091071	AC091071 Oryza sat
c 23	77	8.2	7218	6	I66494	I66494 Sequence 14
24	73.2	7.8	612	8	AF255713	AF255713 Arabidops
25	72	7.7	172224	2	AC087149	AC087149 Mus muscu
26	69.8	7.4	75475	8	AB005235	AB005235 Arabidops
27	69.8	7.4	102897	8	ATF17C15	AL162506 Arabidops
c 28	69.2	7.4	440	8	PAB296729	AJ296729 Picea abi
c 29	68.6	7.3	388	8	PAB296740	AJ296740 Picea abi
c 30	68.2	7.3	20717	2	AC013104	AC013104 Drosophil
31	68.2	7.3	182623	3	AC011071	AC011071 Drosophil
c 32	68.2	7.3	202008	2	AL611970	AL611970 Mus muscu
c 33	68.2	7.3	205429	2	AC005506	AC005506 Plasmodiu
34	68.2	7.3	301457	3	AE003512	AE003512 Drosophil
c 35	67.4	7.2	3720	14	S76368	S76368 ORF 5' of E
c 36	67.4	7.2	43658	14	HSV3PRGEN	M86409 Herpesvirus
c 37	67.4	7.2	112930	14	HSGEND	X64346 Herpesvirus
c 38	66.6	7.1	153477	2	AC006278	AC006278 Plasmodiu
c 39	66.2	7.1	67970	3	PFMAL1P3	AL031746 Plasmodiu
40	66	7.0	397	11	G37798	G37798 GARP Plasmo
41	66	7.0	2384	3	PFAGAR	J03998 Plasmodium
42	66	7.0	158548	3	PFMAL3P2	AL034558 Plasmodiu
43	64.8	6.9	2306	3	DDV18	X15382 D.discoideu
44	64.6	6.9	2486	8	AF026917	AF026917 Zea mays
45	64.6	6.9	179595	2	AC098110	AC098110 Rattus no

ALIGNMENTS

RESULT 1
AX139018
LOCUS AX139018 939 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 5 from Patent EP1094112.
ACCESSION AX139018
VERSION AX139018.1 GI:14274700
KEYWORDS .
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 939)
AUTHORS Wu,K., Miki,B.L., Tian,L. and Brown,D.C.
TITLE Repressing gene expression in plants
JOURNAL Patent: EP 1094112-A 5 25-APR-2001;
The Minister of Agriculture and Agri-Food (CA)
FEATURES Location/Qualifiers
source 1. .939
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
BASE COUNT 296 a 175 c 236 g 232 t
ORIGIN

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 17:56:40 ; Search time 665.94 Seconds
(without alignments)
2420.913 Million cell updates/sec

Title: US-09-645-337-5
Perfect score: 939
Sequence: 1 cacgcgtccgtaaaaatcct.....aaaaaaaaaagggcggccgc 939

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1980.DAT:*
2: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1981.DAT:*
3: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1982.DAT:*
4: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1983.DAT:*
5: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1984.DAT:*
6: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1985.DAT:*
7: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1986.DAT:*
8: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1987.DAT:*
9: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1988.DAT:*
10: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1989.DAT:*
11: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1990.DAT:*
12: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1991.DAT:*
13: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1992.DAT:*
14: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1993.DAT:*
15: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1994.DAT:*
16: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1995.DAT:*
17: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:*
18: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:*
19: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:*
20: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:*
21: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:*
22: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:*
23: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:*
24: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	939	100.0	939	22	AAF80352	Nucleotide sequenc
2	905.4	96.4	991	21	AAC45610	Arabidopsis thalia
3	886	94.4	1002	21	AAC35217	Arabidopsis thalia
4	860	91.6	940	21	AAC36931	Arabidopsis thalia
5	734.4	78.2	774	21	AAC43043	Arabidopsis thalia
6	156.2	16.6	1147	21	AAC41072	Arabidopsis sequenc
7	156.2	16.6	1212	22	AAF80353	Nucleotide sequenc
8	124.2	13.2	741	21	AAC40930	Arabidopsis thalia
9	123.8	13.2	1307	20	AAX90845	Maize histone deac
10	104.2	11.1	1245	20	AAX90844	Maize histone deac
11	100	10.6	1191	20	AAX90843	Maize histone deac
12	98.4	10.5	1283	20	AAX90842	Maize histone deac
13	82.6	8.8	856	21	AAC49668	Arabidopsis thalia
14	77.8	8.3	879	21	AAC34537	Arabidopsis thalia
15	70	7.5	468	21	AAC37832	Arabidopsis thalia
16	68.2	7.3	303	23	ABL24559	Drosophila melanog
c 17	68.2	7.3	2403	23	ABL24558	Drosophila melanog
18	61.6	6.6	4000	18	AAT91902	Mannose-1-phosphat
19	61	6.5	895	21	AAC33860	Arabidopsis thalia
20	60.4	6.4	1260	24	ABI99204	Mouse ischaemic co
21	60	6.4	534	21	AAC39224	Arabidopsis thalia
c 22	59	6.3	475	22	ABA42952	Human breast cell
c 23	59	6.3	475	22	ABA53373	Human foetal liver
c 24	59	6.3	475	22	ABA23153	Probe #1619 for ge
c 25	59	6.3	475	22	AAK01642	Human brain expres
c 26	59	6.3	475	22	AAK27085	Human bone marrow
c 27	59	6.3	475	22	AAI11680	Probe #1613 for ge
c 28	59	6.3	475	22	AAI32980	Probe #1666 used t
c 29	59	6.3	475	22	AAI01605	Probe #1596 used t
30	59	6.3	3224	18	AAT91872	Rat YT521 gene. R
31	58.8	6.3	520	21	AAC43629	Arabidopsis thalia
32	57.8	6.2	2518	20	AAV71745	Human V3 loop HIV
33	57.8	6.2	2518	22	AAC92553	Human nucleolin DN
34	56.8	6.0	3399	17	AAT05868	Chicken leucocytoz
35	56.6	6.0	3211	18	AAT89346	Human p160 cDNA 16
36	56.6	6.0	3901	18	AAT89345	Human p160 cDNA 16
37	56.6	6.0	49999	20	AAZ23891	Murine LOBO genomi
38	56.6	6.0	49999	20	AAZ23896	Murine LOBO homolo
39	56.4	6.0	291	23	AAS68926	DNA encoding novel
40	56.4	6.0	2658	21	AAA70213	Plasmodium falcipa
41	55	5.9	1301	23	AAS69688	DNA encoding novel
42	54.8	5.8	372	23	AAS75452	DNA encoding novel
43	54.8	5.8	801	23	AAS90729	DNA encoding novel
44	54.8	5.8	1072	23	AAS90738	DNA encoding novel
45	54.6	5.8	1024	21	AAZ97203	Human prostate can

ALIGNMENTS

RESULT 1
AAF80352
ID AAF80352 standard; DNA; 939 BP.
XX
AC AAF80352;
XX
DT 29-JUN-2001 (first entry)
XX
DE Nucleotide sequence of a histone deacetylase designated AtHDA2A.
XX
KW Histone deacetylase; AtHDA2A; RPD3; gene expression; transgenic plant;
KW HDA1; ethylene-responsive phenotype; hypocotyl elongation; ds.
XX
OS Arabidopsis thaliana.
XX
FH Key Location/Qualifiers
FT CDS 49..786
FT /*tag= a
FT /product= "histone deacetylase AtHDA2A"
XX
PN CA2316036-A1.
XX
PD 27-FEB-2001.
XX
PF 24-AUG-2000; 2000CA-2316036.
XX
PR 27-AUG-1999; 99US-0383971.
XX
PA (MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.
XX
PI Miki B, Brown D, Tian L, Wu K;
XX

```
DR   WPI; 2001-258457/27.
DR   P-PSDB; AAB67813.
XX
PT   Methods for regulating gene expression in transgenic plants, e.g.
PT   repressing ethylene-responsive phenotypes (e.g. inhibition of hypocotyl
PT   elongation), comprises introducing genes encoding histone deacetylase
PT   -
XX
PS   Claim 9; Fig 2A; 91pp; English.
XX
CC   The present sequence encodes an Arabidopsis thaliana histone deacetylase
CC   designated AtHDA2A. The protein is homologous to yeast RPD3 and HDA1.
CC   The polynucleotide sequence is used in the method of the invention.
CC   The specification describes a method for regulating gene expression in
CC   transgenic plants. The method comprises modifying histones by introducing
CC   chimeric nucleotide sequences which have regulatory elements in operative
CC   association with a gene of interest or with a nucleotide sequence
CC   encoding histone deacetylase. The method is useful for regulating the
CC   developmental, physiological or biochemical pathway within a plant,
CC   particularly for repressing ethylene-responsive phenotypes
CC   (e.g. inhibition of hypocotyl elongation). The method is also useful
CC   as a functional test for identifying a phenotype associated with
CC   perturbing a gene. The histone deacetylase genes are useful for
CC   altering the development of an organism.
XX
SQ   Sequence 939 BP; 296 A; 175 C; 236 G; 232 T; 0 other;


  Query Match             100.0%;  Score 939;  DB 22;  Length 939;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-233;
  Matches  939;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 cacgcgtccgtaaaaatcctctcttttttctcaaccttgattcttagccatggagttctgg 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 cacgcgtccgtaaaaatcctctcttttttctcaaccttgattcttagccatggagttctgg 60

Qy      61 ggaattgaagttaaatcaggaaagccagttacagtgactcctgaagaaggcattcttatc 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ggaattgaagttaaatcaggaaagccagttacagtgactcctgaagaaggcattcttatc 120

Qy     121 cacgtttctcaggcatcgcttggagaatgtaaaaacaagaagggagagtttgtgccttta 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 cacgtttctcaggcatcgcttggagaatgtaaaaacaagaagggagagtttgtgccttta 180

Qy     181 catgtaaaggttgggaaccagaacttggttctgggaactctatcgactgagaacatccct 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 catgtaaaggttgggaaccagaacttggttctgggaactctatcgactgagaacatccct 240

Qy     241 cagcttttctgtgatttggtattcgacaaggagtttgagctttctcacacttggggaaaa 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 cagcttttctgtgatttggtattcgacaaggagtttgagctttctcacacttggggaaaa 300

Qy     301 ggaagtgtttacttttgttggatacaaaactcccaacattgagccacaaggctattctgag 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ggaagtgtttacttttgttggatacaaaactcccaacattgagccacaaggctattctgag 360

Qy     361 gaagaagaggaagaagaggaagaagttcctgctgggaatgctgccaaggctgtagctaaa 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 gaagaagaggaagaagaggaagaagttcctgctgggaatgctgccaaggctgtagctaaa 420

Qy     421 ccaaaggctaagcctgcagaagtgaagccagctgttgatgatgaagaggatgagtctgat 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 ccaaaggctaagcctgcagaagtgaagccagctgttgatgatgaagaggatgagtctgat 480

Qy     481 tctgacggaatggatgaagatgattctgatggtgaggattctgaggaagaagagcctaca 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 tctgacggaatggatgaagatgattctgatggtgaggattctgaggaagaagagcctaca 540

Qy     541 cctaagaagcctgcatcaagcaagaagagagctaatgaaactacccctaaagcacctgtg 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 cctaagaagcctgcatcaagcaagaagagagctaatgaaactacccctaaagcacctgtg 600

Qy     601 tcagcaaagaaggcgaaagtagcagttactcctcagaaaacagatgagaagaagaaaggg 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 tcagcaaagaaggcgaaagtagcagttactcctcagaaaacagatgagaagaagaaaggg 660

Qy     661 ggaaaggctgcaaaccagagcccaaagtcggccagtcaagtctcatgtggttcatgcaag 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 ggaaaggctgcaaaccagagcccaaagtcggccagtcaagtctcatgtggttcatgcaag 720

Qy     721 aagactttcaactcagggaatgcacttgagtctcacaacaaggccaagcacgctgctgcc 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 aagactttcaactcagggaatgcacttgagtctcacaacaaggccaagcacgctgctgcc 780

Qy     781 aagtgaagtggtttcttattagagcttgtgatttctatggaattttgcctgtagtcttta 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 aagtgaagtggtttcttattagagcttgtgatttctatggaattttgcctgtagtcttta 840

Qy     841 tgaaaccttcggattttcttatattttcttttgataacaagagtcttaatgaaagagagc 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 tgaaaccttcggattttcttatattttcttttgataacaagagtcttaatgaaagagagc 900

Qy     901 cagttggagtcttaaaaaaaaaaaaaaaaaaggcggccgc 939
          |||||||||||||||||||||||||||||||||||||||
Db     901 cagttggagtcttaaaaaaaaaaaaaaaaaaggcggccgc 939


RESULT   2
AAC45610
ID   AAC45610 standard; DNA; 991 BP.
XX
AC   AAC45610;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 47120.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
```

PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

```
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match             96.4%;  Score 905.4;  DB 21;  Length 991;
  Best Local Similarity   99.9%;  Pred. No. 8.6e-225;
  Matches  906;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       8 ccgtaaaaatcctctcttttttctcaaccttgattcttagccatggagttctggggaattg 67
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      20 ccctaaaaatcctctctttttctcaaccttgattcttagccatggagttctggggaattg 79

Qy      68 aagttaaatcaggaaagccagttacagtgactcctgaagaaggcattcttatccacgttt 127
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      80 aagttaaatcaggaaagccagttacagtgactcctgaagaaggcattcttatccacgttt 139

Qy     128 ctcaggcatcgcttggagaatgtaaaaacaagaagggagagtttgtgcctttacatgtaa 187
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     140 ctcaggcatcgcttggagaatgtaaaaacaagaagggagagtttgtgcctttacatgtaa 199

Qy     188 aggttgggaaccagaacttggttctgggaactctatcgactgagaacatccctcagcttt 247
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     200 aggttgggaaccagaacttggttctgggaactctatcgactgagaacatccctcagcttt 259

Qy     248 tctgtgatttggtattcgacaaggagtttgagctttctcacacttggggaaaaggaagtg 307
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     260 tctgtgatttggtattcgacaaggagtttgagctttctcacacttggggaaaaggaagtg 319

Qy     308 tttactttgttggatacaaaactcccaacattgagccacaaggctattctgaggaagaag 367
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     320 tttactttgttggatacaaaactcccaacattgagccacaaggctattctgaggaagaag 379

Qy     368 aggaagaagaggaagaagttcctgctgggaatgctgccaaggctgtagctaaaccaaagg 427
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     380 aggaagaagaggaagaagttcctgctgggaatgctgccaaggctgtagctaaaccaaagg 439

Qy     428 ctaagcctgcagaagtgaagccagctgttgatgatgaagaggatgagtctgattctgacg 487
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     440 ctaagcctgcagaagtgaagccagctgttgatgatgaagaggatgagtctgattctgacg 499

Qy     488 gaatggatgaagatgattctgatggtgaggattctgaggaagaagagcctacacctaaga 547
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     500 gaatggatgaagatgattctgatggtgaggattctgaggaagaagagcctacacctaaga 559

Qy     548 agcctgcatcaagcaagaagagagctaatgaaactacccctaaagcacctgtgtcagcaa 607
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     560 agcctgcatcaagcaagaagagagctaatgaaactacccctaaagcacctgtgtcagcaa 619

Qy     608 agaaggcgaaagtagcagttactcctcagaaaacagatgagaagaagaaaggggaaagg 667
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     620 agaaggcgaaagtagcagttactcctcagaaaacagatgagaagaagaaaggggaaagg 679

Qy     668 ctgcaaaccagaagcccaaagtcggccagtcaagtctcatgtggttcatgcaagaagactt 727
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     680 ctgcaaaccagagcccaaagtcggccagtcaagtctcatgtggttcatgcaagaagactt 739

Qy     728 tcaactcagggaatgcacttgagtctcacaacaaggccaagcacgctgctgccaagtgaa 787
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     740 tcaactcagggaatgcacttgagtctcacaacaaggccaagcacgctgctgccaagtgaa 799

Qy     788 gtggtttcttattagagcttgtgatttctatggaattttgcctgtagtctttatgaaacc 847
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     800 gtggtttcttattagagcttgtgatttctatggaattttgcctgtagtctttatgaaacc 859

Qy     848 ttcggatttttcttatatttttcttttgataacaagagtcttaatgaaagagagccagttgg 907
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     860 ttcggatttttcttatatttttcttttgataacaagagtcttaatgaaagagagccagttgg 919

Qy     908 agtctta 914
           |||||||
Db     920 agtctta 926


RESULT   3
AAC35217
ID   AAC35217 standard; DNA; 1002 BP.
XX
AC   AAC35217;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 9400.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
```

PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 94.4%; Score 886; DB 21; Length 1002;
Best Local Similarity 98.8%; Pred. No. 9.2e-220;
Matches 905; Conservative 1; Mismatches 1; Indels 9; Gaps 1;

```
Qy       8 ccgtaaaaatcctctcttttttctcaaccttgattcttagccatggagttctggggaattg 67
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22 ccctaaaaatcctctcttttttctcaaccttgattcttagccatggagttctggggaattg 81

Qy      68 aagttaaatcaggaaagccagttacagtgactcctgaagaaggcattcttatccacgttt 127
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      82 aagttaaatcaggaaagccagttacagtgactcctgaagaaggcattcttatccacgttt 141
```

```
Qy    128 ctcaggcatcgcttggagaatgtaaaaacaagaagggagagtttgtgcctttacatgtaa 187
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    142 ctcaggcatcgcttggagaatgtaaaaacaagaagggagagtttgtgcctttacatgtaa 201

Qy    188 aggttgggaaccagaacttggttctgggaactctatcgactgagaacatccctcagcttt 247
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    202 aggttgggaaccagaacttggttctgggaactctatcgactgagaacatccctcagcttt 261

Qy    248 tctgtgatttggtattcgacaaggagtttgagctttctcacacttggggaaaaggaagtg 307
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    262 tctgtgatttggtattcgacaaggagtttgagctttctcacacttggggaaaaggaagtg 321

Qy    308 tttactttgttggatacaaaactcccaacattgagccacaaggctattct---------g 358
          ||||||||||||||||||||||||||||||||||||||||||||||||         |
Db    322 tttactttgttggatacaaaactcccaacattgagccacaaggctattctgaggaagaag 381

Qy    359 aggaagaagaggaagaagaggaagaagttcctgctgggaatgctgccaaggctgtagcta 418
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    382 aggaagaagaggaagaagaggaagaagttcctgctgggaatgctgccaaggctgtagcta 441

Qy    419 aaccaaaggctaagcctgcagaagtgaagccagctgttgatgatgaagaggatgagtctg 478
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    442 aaccaaaggctaagcctgcagaagtgaagccagctgttgatgatgaagaggatgagtctg 501

Qy    479 attctgacggaatggatgaagatgattctgatggtgaggattctgaggaagaagagcctа 538
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    502 attctgacggaatggatgaagatgattctgatggtgaggattctgaggaagaagagcctа 561

Qy    539 cacctaagaagcctgcatcaagcaagaagagagctaatgaaactacccctaaagcacctg 598
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    562 cacctaagaagcctgcatcaagcaagaagagagctaatgaaactacccctaaagcacctg 621

Qy    599 tgtcagcaaagaaggcgaaagtagcagttactcctcagaaaacagatgagaagaagaaag 658
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    622 tgtcagcaaagaaggcgaaagtagcagttactcctcagaaaacagatgagaagaagaaag 681

Qy    659 ggggaaaggctgcaaaccagagcccaaagtcggccagtcaagtctcatgtggttcatgca 718
          ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db    682 ggggaaaggctgcaaaccakagcccaaagtcggccagtcaagtctcatgtggttcatgca 741

Qy    719 agaagactttcaactcagggaatgcacttgagtctcacaacaaggccaagcacgctgctg 778
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    742 agaagactttcaactcagggaatgcacttgagtctcacaacaaggccaagcacgctgctg 801

Qy    779 ccaagtgaagtggtttcttattagagcttgtgatttctatggaattttgcctgtagtctt 838
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    802 ccaagtgaagtggtttcttattagagcttgtgatttctatggaattttgcctgtagtctt 861

Qy    839 tatgaaaccttcggattttcttatattttcttttgataacaagagtcttaatgaaagaga 898
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    862 tatgaaaccttcggattttcttatattttcttttgataacaagagtcttaatgaaagaga 921

Qy    899 gccagttggagtctta 914
          ||||||||||||||||
Db    922 gccagttggagtctta 937
```

```
RESULT   4
AAC36931
ID   AAC36931 standard; DNA; 940 BP.
XX
AC   AAC36931;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 15573.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
```

```
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
```

PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 91.6%; Score 860; DB 21; Length 940;
Best Local Similarity 97.7%; Pred. No. 4.9e-213;
Matches 899; Conservative 0; Mismatches 5; Indels 16; Gaps 2;

```
Qy      8 ccgtaaaaatcctctcttttttctcaaccttgattcttagccatggagttctggggaattg 67
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     22 ccctaaaaatcctctcttttttctcaaccttgattcttagccatggagttctggggaattg 81

Qy     68 aagttaaatcaggaaagccagttacagtgactcctgaagaaggcattcttatccacgttt 127
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     82 aagttaaatcaggaaagccagttacagtgactcctgaagaaggcattcttatccacgttt 141

Qy    128 ctcaggcatcgcttggagaatgtaaaaacaagaagggagagtttgtgcctttacatgtaa 187
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    142 ctcaggcatcgcttggagaatgtaaaaacaagaagggagagtttgtgcctttacatgtaa 201

Qy    188 aggttgggaaccagaacttggttctgggaactctatcgactgagaacatccctcagcttt 247
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    202 aggttgggaaccagaacttggttctgggaactctatcgactgagaacatccctcagcttt 261

Qy    248 tctgtgatttggtattcgacaaggagtttgagctttctcacacttggggaaaaggaagtg 307
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    262 tctgtgatttggtattcgacaaggagttt-agctttctcacacttggggaaaaggaagtg 320

Qy    308 tttactttgttggatacaaaactcccaacattgagccacaaggctattctgaggaagaag 367
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    321 tttactttgttggatacaaaactcccaacattgagccacaaggctattctgaggaagaag 380

Qy    368 aggaagaagaggaagaagttcctgc----------------tgggaatgctgccaaggctg 412
          |||||||||||||||||||||||||                ||||||||||||||||||
Db    381 aggaagaagaggaagaagttcctgctgccgttactgctagtgggaatgctgccaaggctg 440
```

```
Qy    413 tagctaaaccaaaggctaagcctgcagaagtgaagccagctgttgatgatgaagaggatg 472
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    441 tagctaaaccaaaggctaagcctgcagaagtgaagccagctgttgatgatgaagaggatg 500

Qy    473 agtctgattctgacggaatggatgaagatgattctgatggtgaggattctgaggaagaag 532
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    501 agtctgattctgacggaatggatgaagatgattctgatggtgaggattctgaggaagaag 560

Qy    533 agcctacacctaagaagcctgcatcaagcaagaagagagctaatgaaactacccctaaag 592
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    561 agcctacacctaagaagcctgcatcaagcaagaagagagctaatgaaactacccctaaag 620

Qy    593 cacctgtgtcagcaaagaaggcgaaagtagcagttactcctcagaaaacagatgagaaga 652
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    621 cacctgtgtcagcaaagaaggcgaaagtagcagttactcctcagaaaacagatgagaaga 680

Qy    653 agaaaggggaaaggctgcaaaccagagcccaaagtcggccagtcaagtctcatgtggtt 712
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    681 agaaaggtggaaaggctgcaaaccagagcccaaagtcggccagtcaagtctcatgtggtt 740

Qy    713 catgcaagaagactttcaactcagggaatgcacttgagtctcacaacaaggccaagcacg 772
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    741 catgcaagaagactttcaactcagggaatgcacttgagtctcacaacaaggccaagcacg 800

Qy    773 ctgctgccaagtgaagtggtttcttattagagcttgtgatttctatggaattttgcctgt 832
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    801 ctgctgccaagtgaagtggtttcttattagagcttatgatttctatggaattttgcctgt 860

Qy    833 agtctttatgaaaccttcggattttcttatattttctttgataacaagagtcttaatga 892
          |||||||||||||||||| ||||| ||||||||||||||||||||||||||||||||||||||
Db    861 agtctttatgaaaccttcagatttccttatattttctttgataacaagagtcttaatga 920

Qy    893 aagagagccagttggagtct 912
          ||||||||||||||||||||
Db    921 aagagagccagttggagtct 940
```

RESULT 5
AAC43043
ID AAC43043 standard; DNA; 774 BP.
XX
AC AAC43043;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 37816.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 78.2%; Score 734.4; DB 21; Length 774;
Best Local Similarity 99.9%; Pred. No. 1.6e-180;
Matches 735; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy     49 atggagttctggggaattgaagttaaatcaggaaagccagttacagtgactcctgaagaa 108
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 atggagttctggggaattgaagttaaatcaggaaagccagttacagtgactcctgaagaa 60

Qy    109 ggcattcttatccacgtttctcaggcatcgcttggagaatgtaaaaacaagaagggagag 168
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ggcattcttatccacgtttctcaggcatcgcttggagaatgtaaaaacaagaagggagag 120

Qy    169 tttgtgcctttacatgtaaaggttgggaaccagaacttggttctgggaactctatcgact 228
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 tttgtgcctttacatgtaaaggttgggaaccagaacttggttctgggaactctatcgact 180

Qy    229 gagaacatccctcagcttttctgtgatttggtattcgacaaggagtttgagctttctcac 288
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 gagaacatccctcagcttttctgtgatttggtattcgacaaggagtttgagctttctcac 240

Qy    289 acttggggaaaaggaagtgtttacttttgttggatacaaaactcccaacattgagccacaa 348
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 acttggggaaaaggaagtgtttacttttgttggatacaaaactcccaacattgagccacaa 300

Qy    349 ggctattctgaggaagaagaggaagaagaggaagaagttcctgctgggaatgctgccaag 408
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ggctattctgaggaagaagaggaagaagaggaagaagttcctgctgggaatgctgccaag 360

Qy    409 gctgtagctaaaccaaaggctaagcctgcagaagtgaagccagctgttgatgatgaagag 468
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 gctgtagctaaaccaaaggctaagcctgcagaagtgaagccagctgttgatgatgaagag 420

Qy    469 gatgagtctgattctgacggaatggatgaagatgattctgatggtgaggattctgaggaa 528
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 gatgagtctgattctgacggaatggatgaagatgattctgatggtgaggattctgaggaa 480

Qy    529 gaagagcctacacctaagaagcctgcatcaagcaagaagagagctaatgaaactacccct 588
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 gaagagcctacacctaagaagcctgcatcaagcaagaagagagctaatgaaactacccct 540

Qy    589 aaagcacctgtgtcagcaaagaaggcgaaagtagcagttactcctcagaaaacagatgag 648
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 aaagcacctgtgtcagcaaagaaggcgaaagtagcagttactcctcagaaaacagatgag 600

Qy    649 aagaagaaaggggggaaaggctgcaaaccagagcccaaagtcggccagtcaagtctcatgt 708
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 aagaagaaaggggggaaaggctgcaaaccagagcccaaagtcggccagtcaagtctcatgt 660

Qy    709 ggttcatgcaagaagactttcaactcagggaatgcacttgagtctcacaacaaggccaag 768
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ggttcatgcaagaagactttcaactcagggaatgcacttgagtctcacaacaaggccaag 720
```

```
Qy    769 cacgctgctgccaagt 784
          ||||||||||||||| |
Db    721 cacgctgctgccaatt 736


RESULT  6
AAC41072
ID   AAC41072 standard; DNA; 1147 BP.
XX
AC   AAC41072;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 30546.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
PR   02-AUG-1999;   99US-0146386.
PR   02-AUG-1999;   99US-0146388.
PR   02-AUG-1999;   99US-0146389.
PR   03-AUG-1999;   99US-0147038.
PR   04-AUG-1999;   99US-0147204.
PR   04-AUG-1999;   99US-0147302.
PR   05-AUG-1999;   99US-0147192.
PR   05-AUG-1999;   99US-0147260.
PR   06-AUG-1999;   99US-0147303.
PR   06-AUG-1999;   99US-0147416.
PR   09-AUG-1999;   99US-0147493.
PR   09-AUG-1999;   99US-0147935.
PR   10-AUG-1999;   99US-0148171.
PR   11-AUG-1999;   99US-0148319.
PR   12-AUG-1999;   99US-0148341.
PR   13-AUG-1999;   99US-0148565.
PR   13-AUG-1999;   99US-0148684.
```

```
PR   16-AUG-1999;   99US-0149368.
PR   17-AUG-1999;   99US-0149175.
PR   18-AUG-1999;   99US-0149426.
PR   20-AUG-1999;   99US-0149722.
PR   20-AUG-1999;   99US-0149723.
PR   20-AUG-1999;   99US-0149929.
PR   23-AUG-1999;   99US-0149902.
PR   23-AUG-1999;   99US-0149930.
PR   25-AUG-1999;   99US-0150566.
PR   26-AUG-1999;   99US-0150884.
PR   27-AUG-1999;   99US-0151065.
PR   27-AUG-1999;   99US-0151066.
PR   27-AUG-1999;   99US-0151080.
PR   30-AUG-1999;   99US-0151303.
PR   31-AUG-1999;   99US-0151438.
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match           16.6%;  Score 156.2;  DB 21;  Length 1147;
  Best Local Similarity  64.1%;  Pred. No. 1.5e-30;
  Matches  252;  Conservative    0;  Mismatches  138;  Indels    3;  Gaps    1;

Qy      25 ttttctcaaccttgattcttagccatggagttctggggaattgaagttaaatcaggaaag 84
           || ||||  ||| | |    | | ||||||||||||||||| |||  || | | ||  |||
Db      50 ttctctcttcctcgttcaacaacaatggagttctggggagttgcggtgacaccaaaaaac 109

Qy      85 ccagttacagtgactcctgaagaaggcattcttatccacgtttctcaggcatcgcttgga 144
             |   ||  |||||||||||||||||| ||  ||| ||||| |||||||||| || |||
Db     110 gctactaaggtgactcctgaagaagacagccttgtccacatttctcaggcttcactt--- 166

Qy     145 gaatgtaaaaacaagaagggagagtttgtgccctttacatgtaaaggttgggaaccagaac 204
           || || | |   ||    ||||| | ||||  |||   ||| |  |||||       ||
Db     167 gactgcacagtgaaatctggagaatctgtggttttgagtgtgactgttggtggggctaaa 226

Qy     205 ttggttctgggaactctatcgactgagaacatccctcagcttttctgtgatttggtattc 264
             | ||| | ||||| || ||    || ||  |||||||| ||  || |||||||||| ||
Db     227 cttgttattggaacactttcacaagacaagttccctcagattagctttgatttggttttt 286

Qy     265 gacaaggagtttgagctttctcacacttggggaaaaggaagtgtttactttgttggatac 324
           || || ||||||||||||||| ||||   |   |||| || |||| | ||  |||| |||
Db     287 gataaagagtttgagctttcacacagcggtaccaaagcaaatgttcatttcattggctac 346

Qy     325 aaaactcccaacattgagccacaaggctattctgaggaagaagaggaagaagaggaagaa 384
           ||| | |||||||| ||||   | | ||   ||      || || || || | |     |||
Db     347 aaatcccccaacatcgagcaggatgacttcactagttcggatgatgaggatgttcctgaa 406

Qy     385 gttcctgctgggaatgctgccaaggctgtagct 417
           | |  | |||    |||  | |  ||||| ||
Db     407 gctgttcctgctcctgcccctactgctgttact 439


RESULT   7
AAF80353
ID   AAF80353 standard; DNA; 1212 BP.
XX
AC   AAF80353;
XX
DT   29-JUN-2001  (first entry)
XX
DE   Nucleotide sequence of a histone deacetylase designated AtHDA2B.
XX
KW   Histone deacetylase; AtHDA2B; RPD3; gene expression; transgenic plant;
KW   HDA1; ethylene-responsive phenotype; hypocotyl elongation; ds.
XX
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   CDS             61..978
FT                   /*tag= a
FT                   /product= "histone deacetylase AtHDA2B"
XX
PN   CA2316036-A1.
XX
PD   27-FEB-2001.
XX
PF   24-AUG-2000; 2000CA-2316036.
XX
PR   27-AUG-1999;   99US-0383971.
XX
PA   (MIAC ) CANADA MIN AGRIC & AGRI-FOOD CANADA.
XX
PI   Miki B,  Brown D,  Tian L,  Wu K;
XX
DR   WPI; 2001-258457/27.
DR   P-PSDB; AAB67814.
XX
PT   Methods for regulating gene expression in transgenic plants, e.g.
PT   repressing ethylene-responsive phenotypes (e.g. inhibition of hypocotyl
PT   elongation), comprises introducing genes encoding histone deacetylase
PT   -
XX
PS   Claim 9; Fig 2B; 91pp; English.
XX
CC   The present sequence encodes an Arabidopsis thaliana histone deacetylase
CC   designated AtHDA2B. The protein is homologous to yeast RPD3 and HDA1.
CC   The polynucleotide sequence is used in the method of the invention.
CC   The specification describes a method for regulating gene expression in
CC   transgenic plants. The method comprises modifying histones by introducing
CC   chimeric nucleotide sequences which have regulatory elements in operative
CC   association with a gene of interest or with a nucleotide sequence
CC   encoding histone deacetylase. The method is useful for regulating the
CC   developmental, physiological or biochemical pathway within a plant,
```

```
CC   particularly for repressing ethylene-responsive phenotypes
CC   (e.g. inhibition of hypocotyl elongation). The method is also useful
CC   as a functional test for identifying a phenotype associated with
CC   perturbing a gene. The histone deacetylase genes are useful for
CC   altering the development of an organism.
XX
SQ   Sequence 1212 BP; 393 A; 238 C; 293 G; 288 T; 0 other;


  Query Match             16.6%;  Score 156.2;  DB 22;  Length 1212;
  Best Local Similarity   64.1%;  Pred. No. 1.5e-30;
  Matches  252;  Conservative    0;  Mismatches  138;  Indels    3;  Gaps    1;

Qy      25 ttttctcaaccttgattcttagccatggagttctggggaattgaagttaaatcaggaaag 84
           || ||||  ||| | |    | | ||||||||||||||||| |||  || | | ||  |||
Db      37 ttctctcttcctcgttcaacaacaatggagttctggggagttgcggtgacaccaaaaaac 96

Qy      85 ccagttacagtgactcctgaagaaggcattcttatccacgtttctcaggcatcgcttgga 144
             |   ||  |||||||||||||||| ||  ||| ||||| |||||||||| || |||
Db      97 gctactaaggtgactcctgaagaagacagccttgtccacatttctcaggcttcactt--- 153

Qy     145 gaatgtaaaaacaagaaggggagagtttgtgcctttacatgtaaaggttgggaaccagaac 204
           || || | |   ||    |||| | ||||  |||   ||| |  |||||       ||
Db     154 gactgcacagtgaaatctggagaatctgtggttttgagtgtgactgttggtggggctaaa 213

Qy     205 ttggttctgggaactctatcgactgagaacatccctcagcttttctgtgatttggtattc 264
             | ||| | ||||| || ||    || ||  |||||||| ||  || |||||||| ||
Db     214 cttgttattggaacactttcacaagacaagttccctcagattagctttgatttggttttt 273

Qy     265 gacaaggagtttgagctttctcacacttggggaaaaggaagtgtttactttgttggatac 324
           || || ||||||||||||| ||||   |    |||| || |||| | ||  |||| |||
Db     274 gataaagagtttgagctttcacacagcggtaccaaagcaaatgttcatttcattggctac 333

Qy     325 aaaactcccaacattgagccacaaggctattctgaggaagaagaggaagaagaggaagaa 384
           ||| | |||||||| ||||   | | ||   ||      || || || || |     |||
Db     334 aaatcccccaacatcgagcaggatgacttcactagttcggatgatgaggatgttcctgaa 393

Qy     385 gttcctgctgggaatgctgccaaggctgtagct 417
           | |  | |||    |||  | |  |||||  ||
Db     394 gctgttcctgctcctgcccctactgctgttact 426


RESULT   8
AAC40930
ID   AAC40930 standard; DNA; 741 BP.
XX
AC   AAC40930;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 30042.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
```

PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 13.2%; Score 124.2; DB 21; Length 741;
Best Local Similarity 51.1%; Pred. No. 2.4e-22;
Matches 291; Conservative 0; Mismatches 278; Indels 0; Gaps 0;

```
Qy      45 agccatggagttctggggaattgaagttaaatcaggaaagccagttacagtgactcctga 104
           | | ||||||||||||||  ||||||||||    || || ||| ||    |   |||||
Db     114 aacaatggagttctggggtgttgaagttaagaatggcaaaccacttcacctagatcctgg 173

Qy     105 agaaggcattcttatccacgtttctcaggcatcgcttggagaatgtaaaaacaagaaggg 164
             || ||  ||  | ||| | |||||||   |  | |||||  |||| |||||
Db     174 gctagacaggctcgtgcacatctctcaggttgctttgggagagagtaagaacaatgtaac 233

Qy     165 agagtttgtgcctttacatgtaaaggttgggaaccagaacttggttctgggaactctatc 224
           ||     | |   |  | || |  |||||     | ||  |  |  | ||||| |||||
Db     234 cgaaccaatccagctttacgtgactgttggatcggataagcttctcattggaacgctatc 293

Qy     225 gactgagaacatccctcagcttttctgtgatttggtattcgacaaggagtttgagctttc 284
           ||||||  | ||||||||  |     ||  | || || || | | | |||| ||| ||
Db     294 tcatgagaagtttcctcagctgtctacggagattgtcttggaaaggaactttgcgctgtc 353

Qy     285 tcacacttggggaaaaggaagtgtttactttgttggatacaaaactcccaacattgagcc 344
           ||| ||||||   || || || |||| |||   ||| ||||||  |        ||| ||
Db     354 tcatacttggaagaatgggagcgttttcttctctggttacaaagttgatgcatctgatcc 413

Qy     345 acaaggctattctgaggaagaagaggaagaagaggaagaagttcctgctgggaatgctgc 404
           |     |   |  |   || ||  ||   |||| |  | |      |    | |
Db     414 cgagcctgaggatttgattgatgatcaacttgaggctgctggtttcaaagctgctccgaa 473

Qy     405 caaggctgtagctaaaccaaaggctaagcctgcagaagtgaagccagctgttgatgatga 464
           ||||       |    ||   | ||    || | |  |        |   |  | ||
Db     474 atctgctgcgaaacaggtgaactttcagttgccaaatgaagatgtcaaagccaagcaaga 533

Qy     465 agaggatgagtctgattctgacggaatggatgaagatgattctgatggtgaggattctga 524
           || || |     | |  ||| | |     |  ||||||  |||| ||||  | ||||
Db     534 tgatgacgctgacggtagtgaagaagattcttcagatgatgatgatagtgaaaactctgg 593

Qy     525 ggaagaagagcctacacctaagaagcctgcatcaagcaagaagagagctaatgaaactac 584
           || || |||             |   ||| |||   ||  ||  || | |   | |
Db     594 agacgaggaggaagagaaggttactgctgaatctgacagtgaggaagatgactcatcaga 653

Qy     585 ccctaaagcacctgtgtcagcaaagaagg 613
           |    ||| |  ||  ||  || |  |||
Db     654 cgacgaagaagatgactcctcagaagagg 682
```

RESULT 9
AAX90845

```
ID   AAX90845 standard; DNA; 1307 BP.
XX
AC   AAX90845;
XX
DT   13-JAN-2000  (first entry)
XX
DE   Maize histone deacetylase-9 DNA.
XX
KW   Maize histone deacetylase; family 2, ZmHD2; chromatin structure;
KW   RNA polymerase I; ribosomal RNA production; promoter regulator; promoter;
KW   transcription; plant transformation; heterochromatin; disease resistance;
KW   chromatin assembly; gene activity; toxin screening; pathogenicity;
KW   disease response promoter; ds.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   CDS             87..944
FT                   /*tag= a
FT                   /product= "Maize histone deacetylase"
XX
PN   WO9951731-A2.
XX
PD   14-OCT-1999.
XX
PF   02-APR-1999;   99WO-US07370.
XX
PR   03-APR-1998;   98US-0080563.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Baldwin DA,  Briggs SP,  Crane VC;
XX
DR   WPI; 1999-611038/52.
DR   P-PSDB; AAY28805.
XX
PT   New deacetylase genes, used for producing transgenic plants which have
PT   increased disease resistance  -
XX
PS   Claim 1; Page 80-82; 87pp; English.
XX
CC   The present sequence encodes a maize histone deacetylase . This DNA
CC   belongs to family 2, ZmHD2 and appears to affect chromatin structure at
CC   promoters of RNA polymerase I and thus regulate ribosomal RNA production.
CC   The nucleotide sequence can be used to transform plants and increase
CC   disease resistance by altering the levels of heterochromatin, altering
CC   chromatin assembly, and gene activity of the transformed plants.
CC   Additionally, compositions find use in screening for toxins that affect
CC   pathogenicity and in determining which disease response promoters are
CC   regulated by histone deacetylase.
XX
SQ   Sequence 1307 BP; 386 A; 260 C; 308 G; 353 T; 0 other;
```

```
  Query Match           13.2%;  Score 123.8;  DB 20;  Length 1307;
  Best Local Similarity  54.5%;  Pred. No. 3.9e-22;
  Matches  298;  Conservative    0;  Mismatches  237;  Indels   12;  Gaps    2;

Qy      41 tcttagccatggagttctggggaattgaagttaaatcaggaaagccagttacagtgactc 100
           || | || ||||||||||||||    ||||| ||  |||||    | ||| |    |
Db      79 tcgtcgcaatggagttctgqggtgaagaagtgaagccaggagccacggtttcttgcaaag 138

Qy     101 ctgaagaaggcattcttatccacgtttctcaggcatcgcttggagaatgtaaaaacaaga 160
           ||  || ||  |  |||||||| |||| |||||  | || || |||     ||| ||
Db     139 ttggtgatggtttggttatccacctttcacaggctgccctaggggaa---ccaaagaaag 195

Qy     161 aqqqaqaqtttgtgcctttacatgtaaaggttgggaaccaqaacttggttctgggaactc 220
           | | |||  ||    | |   ||| ||  |||   |  ||||  | || || |||||
Db     196 cgagtgagaatgccattgtgtctgtcaaaattgatgataagaaactagtgcttggaacct 255

Qy     221 tatcgactgagaacatccctcagcttttctgtgatttggtattcgacaaggagtttgagc 280
           ||||   ||||||    |||||  | |  |||||| ||||||| || || || ||||||
Db     256 tatcagttgagaagcatcctcaaatctcttgtgatctggtatttgataaagattttgagt 315

Qy     281 tttctcacacttggggaaaaggaagtgtttactttgttggatacaaaactcccaacattg 340
           | || |||| ||     | ||  ||||||| |||   ||| |||||  | ||      |
Db     316 tatcacacaattcaaagacagctagtgttttcttctgtggctacaagtcacctgttcctc 375

Qy     341 agccacaaggctattctgaggaagaagaggaagaagaggaagaagttcctgctgqggaatg 400
           |    |     ||||||  |||||      ||| || || |  |  || | |   | |
Db     376 tgtttgagtctgattctggtgaagatagttcagatgaagaggttgaacccgatctaattc 435

Qy     401 ctgccaaggctgtagctaaaccaaaggcta---------agcctgcagaagtgaagccag 451
           |     ||  |   |  |    ||   |||           || |   | ||      |
Db     436 caatgcagaataatgaaattaaaatttctactgcaaaggttcccgtgaaggttggtatac 495

Qy     452 ctgttgatgatgaaqaggatgaqtctgattctgacggaatggatgaaqatgattctqatg 511
           || |||||||||| ||    |||  |  ||| |     |||     ||||  ||||
Db     496 aaaatgctgatgaagatgaaacttctagtggtgatgatgatgatttcactgatagtgata 555

Qy     512 gtgaggattctgaggaagaagagcctacacctaagaagcctgcatcaagcaagaagagag 571
           ||||    |||||||||||| ||| | |    | |  |    |  ||||||  | |
Db     556 gtgaaaatgtctgaggaagatgagtccagtgatgaagatgaagtgtcaagcgatacagata 615

Qy     572 ctaatga 578
           ||| |||
Db     616 ctagtga 622
```

```
RESULT  10
AAX90844
ID   AAX90844 standard; DNA; 1245 BP.
XX
AC   AAX90844;
XX
DT   13-JAN-2000  (first entry)
XX
DE   Maize histone deacetylase-8 DNA.
XX
KW   Maize histone deacetylase; family 2, ZmHD2; chromatin structure;
KW   RNA polymerase I; ribosomal RNA production; promoter regulator; promoter;
KW   transcription; plant transformation; heterochromatin; disease resistance;
KW   chromatin assembly; gene activity; toxin screening; pathogenicity;
KW   disease response promoter; ds.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   CDS             84..1019
FT                   /*tag= a
FT                   /product= "Maize histone deacetylase"
XX
PN   WO9951731-A2.
XX
PD   14-OCT-1999.
XX
PF   02-APR-1999;   99WO-US07370.
XX
PR   03-APR-1998;   98US-0080563.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Baldwin DA,  Briggs SP,  Crane VC;
XX
DR   WPI; 1999-611038/52.
DR   P-PSDB; AAY28804.
XX
PT   New deacetylase genes, used for producing transgenic plants which have
PT   increased disease resistance  -
XX
PS   Claim 1; Page 77-79; 87pp; English.
XX
CC   The present sequence encodes a maize histone deacetylase . This DNA
CC   belongs to family 2, ZmHD2 and appears to affect chromatin structure at
CC   promoters of RNA polymerase I and thus regulate ribosomal RNA production.
CC   The nucleotide sequence can be used to transform plants and increase
```

CC disease resistance by altering the levels of heterochromatin, altering
CC chromatin assembly, and gene activity of the transformed plants.
CC Additionally, compositions find use in screening for toxins that affect
CC pathogenicity and in determining which disease response promoters are
CC regulated by histone deacetylase.
XX
SQ Sequence 1245 BP; 385 A; 250 C; 324 G; 286 T; 0 other;

Query Match 11.1%; Score 104.2; DB 20; Length 1245;
Best Local Similarity 53.7%; Pred. No. 4.6e-17;
Matches 267; Conservative 0; Mismatches 218; Indels 12; Gaps 2;

```
Qy     47 ccatggagttctggggaattgaagttaaatcaggaaagccagttacagtgactcctgaag 106
Db     82 cgatggagttctggggtctcgaggtcaaacctggatccactgtcaagtgtgagcctggac 141

Qy    107 aaggcattcttatccacgtttctcaggcatcgcttggagaatgtaaaaacaagaagggag 166
Db    142 atggctttatcctgcacgtttcccaggctgcccttggggaatcaaagaaaagtgacagtg 201

Qy    167 agtttgtgcctttacatgtaaaggttgggaaccagaacttggttctgggaactctatcga 226
Db    202 ccttaatg------tatgtcaaagttgatgacaagaagcttgccattggaacgctctcta 255

Qy    227 ctgagaacatccctcagcttttctgtgatttggtattcgacaaggagtttgagctttctc 286
Db    256 tcgacaaatacccacagatacaattcgatttggtttttcaataaagagtttgagctgtcac 315

Qy    287 acacttggggaaaaggaagtgtttacttttgttggatacaaaact------cccaacattg 340
Db    316 acacatcgaaaactaccagtgtatttttctctggttacaaggttgagcagccaattgagg 375

Qy    341 agccacaaggctattctgaggaagaagaggaagaagaggaagaagttcctgctgggaatg 400
Db    376 gagatgaaatggatcttgattctgaggatgaagaggaggagctaaacattccagtaatca 435

Qy    401 ctgccaaggctgtagctaaaccaaaggctaagcctgcagaagtgaagccagctgttgatg 460
Db    436 aggaaaatggcaaagctgatgggaaggaggagcagaaaaatcaagagaaggcagtagctg 495

Qy    461 atgaagaggatgagtctgattctgacggaatggatgaagatgattctgatggtgaggatt 520
Db    496 ctacagcttcaaaatcaagtcttggccttgaaaagaaaagcaaggatgactctgatgatt 555

Qy    521 ctgaggaagaagagcct 537
Db    556 ctgatgaggatgagtct 572
```

RESULT 11
AAX90843
ID AAX90843 standard; DNA; 1191 BP.
XX
AC AAX90843;
XX
DT 13-JAN-2000 (first entry)
XX
DE Maize histone deacetylase-7 DNA.
XX
KW Maize histone deacetylase; family 2, ZmHD2; chromatin structure;
KW RNA polymerase I; ribosomal RNA production; promoter regulator; promoter;
KW transcription; plant transformation; heterochromatin; disease resistance;
KW chromatin assembly; gene activity; toxin screening; pathogenicity;
KW disease response promoter; ds.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 63..971
FT /*tag= a
FT /product= "Maize histone deacetylase"
XX

PN WO9951731-A2.
XX
PD 14-OCT-1999.
XX
PF 02-APR-1999; 99WO-US07370.
XX
PR 03-APR-1998; 98US-0080563.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Baldwin DA, Briggs SP, Crane VC;
XX
DR WPI; 1999-611038/52.
DR P-PSDB; AAY28803.
XX
PT New deacetylase genes, used for producing transgenic plants which have
PT increased disease resistance -
XX
PS Claim 1; Page 73-75; 87pp; English.
XX
CC The present sequence encodes a maize histone deacetylase . This DNA
CC belongs to family 2, ZmHD2 and appears to affect chromatin structure at
CC promoters of RNA polymerase I and thus regulate ribosomal RNA production.
CC The nucleotide sequence can be used to transform plants and increase
CC disease resistance by altering the levels of heterochromatin, altering
CC chromatin assembly, and gene activity of the transformed plants.
CC Additionally, compositions find use in screening for toxins that affect
CC pathogenicity and in determining which disease response promoters are
CC regulated by histone deacetylase.
XX
SQ Sequence 1191 BP; 365 A; 242 C; 311 G; 273 T; 0 other;

Query Match 10.6%; Score 100; DB 20; Length 1191;
Best Local Similarity 54.9%; Pred. No. 5.5e-16;
Matches 268; Conservative 0; Mismatches 205; Indels 15; Gaps 3;

```
Qy     49 atggagttctggggaattgaagttaaatcaggaaagccagttacagtgactcctgaagaa 108
Db     63 atggagttctggggtctcgaggtcaagcctggttccactgttaagtgtgagcctggatat 122

Qy    109 ggcattcttatccacgtttctcaggcatcgcttggagaatgtaaaaacaagaagggagag 168
Db    123 ggctttgtgctgcacctttcccaggctgctcttggggaatcg------aagaagagtgat 176

Qy    169 tttgtgcctttacatgtaaaggttgggaaccagaacttggttctgggaactctatcgact 228
Db    177 aatgccttgatgtatgtcaaaattgatgatcagaaacttgccattggaaccctctctgtt 236

Qy    229 gagaacatccctcagctttctgtgatttggtattcgacaaggagtttgagctttctcac 288
Db    237 gacaagaacccacacattcaatttgatctgattttcgataaagagtttgagctgtcgcac 296

Qy    289 acttggggaaaaggaagtgtttactttgttggataca------aaactcccaacattgag 342
Db    297 acatcaaaaactaccagcgtcttcttcactggctacaaggttgaacagccattcgaggaa 356

Qy    343 ccacaaggctattctgaggaagaagaggaagaagaggaagaagttcctgctgggaatgct 402
Db    357 gatgaaatggatcttgattctgaagatgaagacgaggagctgaatgttccagtagtcaag 416

Qy    403 gccaaggctgtagctaaaccaaaggctaagcctgcagaagtgaagccagctgttgatgat 462
Db    417 gaaaatggcaaagctgatgggaagaaacagaaaagtcaagaaaaggcagttgctgcacct 476

Qy    463 gaagaggatgagtctgattc---tgacggaatggatgaagatgattctgatggtgaggat 519
Db    477 tcaaaatcaagtccggattccaagaagagcaaggatgacgatgattctgatgaggacgag 536

Qy    520 tctgagga 527
Db    537 actgatga 544
```

```
RESULT 12
AAX90842
ID   AAX90842 standard; DNA; 1283 BP.
XX
AC   AAX90842;
XX
DT   13-JAN-2000 (first entry)
XX
DE   Maize histone deacetylase-6 DNA.
XX
KW   Maize histone deacetylase; family 2, ZmHD2; chromatin structure;
KW   RNA polymerase I; ribosomal RNA production; promoter regulator; promoter;
KW   transcription; plant transformation; heterochromatin; disease resistance;
KW   chromatin assembly; gene activity; toxin screening; pathogenicity;
KW   disease response promoter; ds.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   CDS             38..1336
FT                   /*tag= a
FT                   /product= "Maize histone deacetylase"
XX
PN   WO9951731-A2.
XX
PD   14-OCT-1999.
XX
PF   02-APR-1999;   99WO-US07370.
XX
PR   03-APR-1998;   98US-0080563.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Baldwin DA, Briggs SP, Crane VC;
XX
DR   WPI; 1999-611038/52.
DR   P-PSDB; AAY28802.
XX
PT   New deacetylase genes, used for producing transgenic plants which have
PT   increased disease resistance -
XX
PS   Claim 1; Page 70-72; 87pp; English.
XX
CC   The present sequence encodes a maize histone deacetylase . This DNA
CC   belongs to family 2, ZmHD2 and appears to affect chromatin structure at
CC   promoters of RNA polymerase I and thus regulate ribosomal RNA production.
CC   The nucleotide sequence can be used to transform plants and increase
CC   disease resistance by altering the levels of heterochromatin, altering
CC   chromatin assembly, and gene activity of the transformed plants.
CC   Additionally, compositions find use in screening for toxins that affect
CC   pathogenicity and in determining which disease response promoters are
CC   regulated by histone deacetylase.
XX
SQ   Sequence 1283 BP; 383 A; 259 C; 333 G; 308 T; 0 other;

  Query Match            10.5%;  Score 98.4;  DB 20;  Length 1283;
  Best Local Similarity  54.7%;  Pred. No. 1.5e-15;
  Matches  267;  Conservative    0;  Mismatches  206;  Indels   15;  Gaps    3;

Qy     49 atggagttctggggaattgaagttaaatcaggaaagccagttacagtgactcctgaagaa 108
          ||||||||||||||  | || || ||  | ||     | ||||        |||| | |
Db     79 atggagttctggggtctcgaggtcaagcctggttccactgttaagtgtgagcctggatat 138

Qy    109 ggcattcttatccacgtttctcaggcatcgcttggagaatgtaaaaacaagaagggagag 168
          ||| || |  | ||| |||| |||||  | ||||| ||||       ||||||  | ||
Db    139 ggctttgtgctgcacctttcccaggctgctcttggggaatcg------aaqaagagtgat 192

Qy    169 tttgtgccttacatgtaaaggttgggaaccagaacttggttctgggaactctatcgact 228
            ||      |  |||| ||  |||   | |||||  | |   | |||||| ||  |
Db    193 aatgccttgatgtatgtcaaaattgatgatcagaaacttgccattggaaccctctctgtt 252

Qy    229 gagaacatccctcagcttttctgtgatttggtattcgacaaggagtttgagctttctcac 288
          || || | ||| ||  |    | |||| || | ||||| || ||||||||||| || |||
Db    253 gacaagaacccacacatacaatttgatctgattttcgataaagaqtttgagctgtcgcac 312

Qy    289 acttggggaaaaggaagtgtttacttttgttggatacaaaact------cccaacattgaq 342
          || |    ||     || |||| |||   ||| |||||   |      ||   |  ||
Db    313 acatcaaaaactaccagcgttttcttcactggctacaaggttgagcagccattcgaggaa 372

Qy    343 ccacaaggctattctgaggaagaagaggaagaagaggaagaagttcctgctgggaatgct 402
              ||    ||  |||    ||||| ||||| |||||       |  | | |
Db    373 gatgaaatggatcttgattctgaagatgaagacgaggagctgaatgttccagtagtcaaq 432

Qy    403 gccaaggctgtagctaaaccaaaggctaagcctgcagaagtgaagccagctgttgatgat 462
          |  || |    |||| |    |||    ||       |||  ||| ||| || ||    |
Db    433 gaaaatggcaaagctgatgaqaagaaacagaaaaqtcaagaaaaggcagttgctgcacct 492

Qy    463 gaagaggatgagtctgattc---tgacggaatggatgaagatgattctgatggtgaggat 519
            | |        | |||||     |  | | |||||| || ||||||||||  || ||
Db    493 tcaaaatcaagtccggattccaagaagagcaaggatgacgacgattctgatgaggacgag 552

Qy    520 tctgagga 527
            |||| ||
Db    553 actgatga 560


RESULT 13
AAC49668
ID   AAC49668 standard; DNA; 856 BP.
XX
AC   AAC49668;
XX
DT   18-OCT-2000 (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 61995.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
```

PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.

PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 8.8%; Score 82.6; DB 21; Length 856;
Best Local Similarity 53.0%; Pred. No. 1.6e-11;
Matches 271; Conservative 0; Mismatches 229; Indels 11; Gaps 4;

```
Qy      30 tcaaccttgattcttagccatggagttctggggaattgaagttaaatcaggaaagccagt 89
           ||  |||   |   |||| |||||||| ||||| || ||  ||||  |||| |||||| |
Db      71 tcttcctctttcactagctatggagttttggggtatcgagattaagccagggaagccatt 130

Qy      90 tacagtgactcctgaagaaggcattcttatccacgtttctcaggcatcgcttggagaatg 149
           ||  ||||  |   |||| ||  |  |  |||| |  |||||||   | ||||| ||
Db     131 taaggtgatacaaaaagatggattcatggtccatgcctctcaggttacccttggtgacgt 190

Qy     150 taaaaa---caagaagggagagtttgtgcctttacatgtaaaggttgg---gaaccagaa 203
           | | ||    || || |  |||  | |  |  |  |||| ||| ||||     |  ||||
Db     191 tgagaaggttaaaaaagatgagacttttgccgtttatgtgaagattggtgatgatgagaa 250

Qy     204 cttggttctgggaactctatcgactgagaacatccctcagcttttctgtgatttggtatt 263
             | || ||             |  ||||  | |||||  |||    |||| |    ||
Db     251 tgggtttatgattggaaatctctcacagaagtttcctcaattttctattgatctctactt 310

Qy     264 cgacaaggagtttgagctttctcacacttggggaaaaggaagtgtttactttgttggata 323
           |   | |||||||||| |||||||||       | | | | |||||  ||  || |||| ||
Db     311 agggcacgagtttgagatttctcaca---acagtacaagcagtgtctatcttattggtta 367

Qy     324 caaaactcccaacattgagccacaaggctattctgaggaagaagaggaagaagaggaaga 383
           ||  ||      |   |       ||  ||||||||||  |||  ||  |   ||  || |
Db     368 caggacctttgatgcttttgacgaa--ctattctgagttagatgaatatatggaacaaca 425

Qy     384 agttcctgctgggaatgctgccaaggctgtagctaaaccaaaggctaagcctgcagaagt 443
           | || |||||  |  |       ||       || ||  | | | | |  | |    ||
Db     426 aattgctgctttgcctcaaaatgagatcaatcctgaagaagatgatgaatccgactcaga 485

Qy     444 gaagccagctgttgatgatgaagaggatgagtctgattctgacggaatggatgaagatga 503
             ||   | | | || || || || ||    || |||   || |   | || |  || |
Db     486 tgagatgggtttggacgaggatgatgactcttcagatgaagaagatgtagaggctgaagc 545

Qy     504 ttctgatggtgaggattctgaggaagaagag 534
             ||      | || | ||  |    || ||
Db     546 acctttaaaggtggctcctccgagcaaaaag 576
```

RESULT 14
AAC34537
ID AAC34537 standard; DNA; 879 BP.
XX
AC AAC34537;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 6997.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 8.3%; Score 77.8; DB 21; Length 879;
Best Local Similarity 52.0%; Pred. No. 2.8e-10;
Matches 275; Conservative 0; Mismatches 242; Indels 12; Gaps 4;

```
Qy     30 tcaaccttgattcttagccatggagttctggggaattgaagttaaatcaggaaagccagt 89
          ||  |||    |   |||| |||||||| ||||| || ||  ||||  |||| |||||| |
Db     73 tcttcctctttcactagctatggagttttggggtatcgagattaagccagggaagccatt 132

Qy     90 tacagtgactcctgaagaaggcattcttatccacgtttctcaggcatcgcttggagaatg 149
          ||  ||||  |   |||| ||  |  |  |||| |  |||||||   | |||||  ||
Db    133 taaggtgatacaaaaagatggattcatggtccatgcctctcaggttacccttggtgacgt 192

Qy    150 taaaaa---caagaagggagagtttgtgcctttacatgtaaaggttgg---gaaccagaa 203
          | | ||    || || |  |||  | |  |  |  |||| ||| ||||     |  ||||
Db    193 tgagaaggttaaaaaagatgagacttttgccgtttatgtgaagattggtgatgatgagaa 252

Qy    204 cttggttctgggaactctatcgactgagaacatccctcagcttttctgtgatttggtatt 263
          | || ||             |  ||||  | |||||  |||    ||||  |    ||
Db    253 tgggtttatgattggaaatctctcacagaagtttcctcaattttctattgatctctactt 312

Qy    264 cgacaaggagtttgagctttctcacacttggggaaaaggaagtgtttacttttgttggata 323
          |   | |||||||||| ||||||||    |    | | | ||||| ||  || |||| ||
Db    313 agggcacgagtttgagatttctcacaacagt---acaagcagtgtctatcttattggtta 369

Qy    324 caaaactcccaacattgagccacaaggctattctgaggaagaagaggaagaagaggaaga 383
          ||  ||      |   |       ||    ||   |||   ||    ||    ||| |||
Db    370 caggacctttgatgcttttgacgaactggatgaggagattgattctgattctgagttaga 429

Qy    384 agttcctgctgggaatgctgccaaggctgtagctaaaccaaaggctaagcctgcagaagt 443
          |    |   |   |           ||| |  || ||    ||   || |||| |||||
Db    430 tgaatatatggaacaacaaattgctgctttgcctcaaaatgagatcaatcctgaagaaga 489

Qy    444 gaagccagctgttgatgatgaagaggat---gagtctgattctgacggaatggatgaaga 500
          |   | | |     |||||   || |   ||    |||  ||||      ||||||||
Db    490 tgatgaatccgactcagatgagatgggtttggacgaggatgatgactcttcagatgaaga 549

Qy    501 tgattctgatggtgaggattctgaggaagaagagcctacacctaagaag 549
          |||   || | ||| |   ||    | |  |  ||| |    || |||
Db    550 agatgtagaggctgaagcacctttaaaggtggctcctccgagcaaaaag 598
```

```
RESULT 15
AAC37832
ID   AAC37832 standard; DNA; 468 BP.
XX
AC   AAC37832;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 18825.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
PR   02-AUG-1999;   99US-0146386.
PR   02-AUG-1999;   99US-0146388.
PR   02-AUG-1999;   99US-0146389.
PR   03-AUG-1999;   99US-0147038.
PR   04-AUG-1999;   99US-0147204.
PR   04-AUG-1999;   99US-0147302.
PR   05-AUG-1999;   99US-0147192.
PR   05-AUG-1999;   99US-0147260.
PR   06-AUG-1999;   99US-0147303.
PR   06-AUG-1999;   99US-0147416.
PR   09-AUG-1999;   99US-0147493.
PR   09-AUG-1999;   99US-0147935.
PR   10-AUG-1999;   99US-0148171.
PR   11-AUG-1999;   99US-0148319.
PR   12-AUG-1999;   99US-0148341.
PR   13-AUG-1999;   99US-0148565.
PR   13-AUG-1999;   99US-0148684.
PR   16-AUG-1999;   99US-0149368.
PR   17-AUG-1999;   99US-0149175.
PR   18-AUG-1999;   99US-0149426.
```

PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 7.5%; Score 70; DB 21; Length 468;
Best Local Similarity 57.4%; Pred. No. 2.3e-08;
Matches 124; Conservative 1; Mismatches 91; Indels 0; Gaps 0;

Qy 437 cagaagtgaaqccagctgttgatgatgaagaggatgagtctgattctgacggaatggatg 496
Db 252 ctgaagaggatgatgactctgagaaaggaatggatgttgatgaagatgactcagatgatg 311

Qy 497 aagatgattctgatggtgaggattctgaggaagaagagcctacacctaagaagcctgcat 556
Db 312 acgaggaggaggattctgaggatgaagaagaggaggagactcctaataagcctgagccaa 371

Qy 557 caagcaagaagagagctaatgaaactacccctaaagcacctgtgtcagcaaagaaggcga 616

Db 372 tcaacaagaagaggccaaatgaatctgtatccaraacacccgtctctggaaagaaggcaa 431

Qy 617 aagtagcagttactcctcagaaaacagatgagaaga 652
Db 432 aaccagcagcagcaccagcttctactcctcagaaga 467

Search completed: May 2, 2002, 20:26:19
Job time: 8979 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 16:03:33 ; Search time 100.71 Seconds
(without alignments)
2290.238 Million cell updates/sec

Title: US-09-645-337-5
Perfect score: 939
Sequence: 1 cacgcgtccgtaaaaatcct.....aaaaaaaaaagggcggccgc 939

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	123.8	13.2	1307	4	US-09-282-305-17	Sequence 17, Appl
2	104.2	11.1	1245	4	US-09-282-305-15	Sequence 15, Appl
3	100	10.6	1191	4	US-09-282-305-13	Sequence 13, Appl
4	98.4	10.5	1283	4	US-09-282-305-11	Sequence 11, Appl
c 5	77	8.2	7218	1	US-08-232-463-14	Sequence 14, Appl
6	57.8	6.2	2518	4	US-09-433-699-3	Sequence 3, Appli
7	56.6	6.0	3211	2	US-08-574-959A-8	Sequence 8, Appli
8	56.6	6.0	3211	4	US-09-357-014-8	Sequence 8, Appli
9	56.6	6.0	3901	2	US-08-574-959A-6	Sequence 6, Appli
10	56.6	6.0	3901	4	US-09-357-014-6	Sequence 6, Appli
11	55.6	5.9	1164	4	US-09-538-871-3	Sequence 3, Appli
12	55.6	5.9	1232	4	US-09-538-871-1	Sequence 1, Appli
13	52.2	5.6	3489	2	US-08-728-323A-1	Sequence 1, Appli
c 14	52.2	5.6	32207	2	US-08-770-379-20	Sequence 20, Appl
c 15	52.2	5.6	32207	4	US-08-757-669A-20	Sequence 20, Appl
c 16	52.2	5.6	32207	4	US-09-230-371A-20	Sequence 20, Appl
17	49.4	5.3	2295	1	US-08-375-300-3	Sequence 3, Appli
18	49.4	5.3	2295	3	US-09-177-431-3	Sequence 3, Appli
19	49.4	5.3	2295	5	PCT-US95-16930-3	Sequence 3, Appli
20	49.4	5.3	4080	1	US-08-375-300-1	Sequence 1, Appli
21	49.4	5.3	4080	3	US-09-177-431-1	Sequence 1, Appli
22	49.4	5.3	4080	5	PCT-US95-16930-1	Sequence 1, Appli
23	48.6	5.2	966	2	US-08-766-738-2	Sequence 2, Appli
24	48.4	5.2	759	1	US-08-466-603-4	Sequence 4, Appli
25	48.4	5.2	759	1	US-08-314-503A-4	Sequence 4, Appli
26	48.4	5.2	759	1	US-08-468-066-4	Sequence 4, Appli
27	48.4	5.2	759	2	US-08-466-717-4	Sequence 4, Appli
28	48.4	5.2	759	3	US-08-466-743-4	Sequence 4, Appli
29	48.4	5.2	759	5	PCT-US95-12414-4	Sequence 4, Appli
30	48.4	5.2	980	1	US-08-466-603-3	Sequence 3, Appli
31	48.4	5.2	980	1	US-08-314-503A-3	Sequence 3, Appli
32	48.4	5.2	980	1	US-08-468-066-3	Sequence 3, Appli
33	48.4	5.2	980	2	US-08-466-717-3	Sequence 3, Appli
34	48.4	5.2	980	3	US-08-466-743-3	Sequence 3, Appli
35	48.4	5.2	980	5	PCT-US95-12414-3	Sequence 3, Appli
36	47	5.0	289	4	US-09-007-005-17	Sequence 17, Appl
37	47	5.0	289	4	US-09-244-796-17	Sequence 17, Appl
38	47	5.0	2277	1	US-08-676-967-2	Sequence 2, Appli
39	47	5.0	2277	1	US-08-676-974-2	Sequence 2, Appli
40	47	5.0	2277	2	US-09-098-487-2	Sequence 2, Appli
41	46.4	4.9	2787	4	US-09-277-565-27	Sequence 27, Appl
42	46	4.9	1276	4	US-09-177-325-2	Sequence 2, Appli
43	46	4.9	1276	4	US-09-411-812A-2	Sequence 2, Appli
44	46	4.9	1276	4	US-09-590-113-2	Sequence 2, Appli
45	45.6	4.9	43795	3	US-08-742-185-101	Sequence 101, App

ALIGNMENTS

RESULT 1
US-09-282-305-17
; Sequence 17, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
; APPLICANT: Baldwin, Donald A.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Crane, Virginia C.
; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
; FILE REFERENCE: 5718-44,
; CURRENT APPLICATION NUMBER: US/09/282,305
; CURRENT FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 60/080,563
; PRIOR FILING DATE: 1998-04-03
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 1307
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (87)..(944)
US-09-282-305-17

Query Match 13.2%; Score 123.8; DB 4; Length 1307;
Best Local Similarity 54.5%; Pred. No. 1e-24;
Matches 298; Conservative 0; Mismatches 237; Indels 12; Gaps 2;

```
Qy      41 tcttagccatggagttctggggaattgaagttaaatcaggaaagccagttacagtgactc 100
           || | || ||||||||||||||||    ||||| ||  |||||    | ||| |    |
Db      79 tcgtcgcaatggagttctggggtgaagaagtgaagccaggagccacggtttcttgcaaag 138

Qy     101 ctgaagaaggcattcttatccacgtttctcaggcatcgcttggagaatgtaaaaacaaga 160
             ||  || ||  |  |||||||| |||| |||||  | || || |||     ||| ||
Db     139 ttggtgatggtttggttatccaccttcacaggctgccctaggggaa---ccaaagaaag 195

Qy     161 agggagagtttgtgcctttacatgtaaaggttgggaaccagaacttggttctgggaactc 220
             | | |||  ||    | |   ||| ||  |||   |  ||||  | || || |||||
Db     196 cgagtgagaatgccattgtgtctgtcaaaattgatgataagaaactagtgcttggaacct 255

Qy     221 tatcgactgagaacatccctcagcttttctgtgatttggtattcgacaaggagtttgagc 280
           ||||   ||||||    |||||  | |  |||||| ||||||| || || || ||||||
Db     256 tatcagttgagaagcatcctcaaatctcttgtgatctggtatttgataaagattttgagt 315

Qy     281 tttctcacacttggggaaaaggaagtgtttacttttgttggatacaaaactcccaacattg 340
           | || |||| ||     | ||  |||||||| |||   ||| |||||  | ||      |
Db     316 tatcacacaattcaaagacagctagtgttttcttctgtggctacaagtcacctgttcctc 375
```

```
Qy    341 agccacaaggctattctgaggaagaagaggaagaagaggaagaagttcctgctgggaatg 400
Db    376 tgtttgagtctgattctggtgaagatagttcagatgaagaggttgaacccgatctaattc 435

Qy    401 ctgccaaggctgtagctaaaccaaaggcta---------agcctgcagaagtgaagccag 451
Db    436 caatgcagaataatgaaattaaaatttctactgcaaaggttcccgtgaaggttggtatac 495

Qy    452 ctgttgatgatgaagaggatgagtctgattctgacggaatggatgaagatgattctgatg 511
Db    496 aaaatgctgatgaagatgaaacttctagtggtgatgatgatgatttcactgatagtgata 555

Qy    512 gtgaggattctgaggaagaagagcctacacctaagaagcctgcatcaagcaagaagagag 571
Db    556 gtgaaatgtctgaggaagatgagtccagtgatgaagatgaagtgtcaagcgatacagata 615

Qy    572 ctaatga 578
Db    616 ctagtga 622
```

RESULT 2
US-09-282-305-15
; Sequence 15, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
; APPLICANT: Baldwin, Donald A.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Crane, Virginia C.
; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
; FILE REFERENCE: 5718-44,
; CURRENT APPLICATION NUMBER: US/09/282,305
; CURRENT FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 60/080,563
; PRIOR FILING DATE: 1998-04-03
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 1245
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (84)..(1019)
US-09-282-305-15

Query Match 11.1%; Score 104.2; DB 4; Length 1245;
Best Local Similarity 53.7%; Pred. No. 2.4e-19;
Matches 267; Conservative 0; Mismatches 218; Indels 12; Gaps 2;

```
Qy     47 ccatggagttctggggaattgaagttaaatcaggaaagccagttacagtgactcctgaag 106
Db     82 cgatggagttctggggtctcgaggtcaaacctggatccactgtcaagtgtgagcctggac 141

Qy    107 aaggcattcttatccacgtttctcaggcatcgcttggagaatgtaaaaacaagaagggag 166
Db    142 atggctttatcctgcacgtttcccaggctgcccttggggaatcaaagaaaagtgacagtg 201

Qy    167 agtttgtgcctttacatgtaaaggttgggaaccagaacttggttctgggaactctatcga 226
Db    202 ccttaatg------tatgtcaaagttgatgacaagaagcttgccattggaacgctctcta 255

Qy    227 ctgagaacatccctcagcttttctgtgatttggtattcgacaaggagtttgagctttctc 286
Db    256 tcgacaaatacccacagatacaattcgatttggttttcaataaagagtttgagctgtcac 315

Qy    287 acacttggggaaaaggaagtgtttactttgttggatacaaaact------cccaacattg 340
Db    316 acacatcgaaaactaccagtgtatttttctctggttacaaggttgagcagccaattgagg 375

Qy    341 agccacaaggctattctgaggaagaagaggaagaagaggaagaagttcctgctgggaatg 400
```

```
Db    376 gagatgaaatggatcttgattctgaggatgaagaggaggagctaaacattccagtaatca 435

Qy    401 ctgccaaggctgtagctaaaccaaaggctaagcctgcagaagtgaagccagctgttgatg 460
Db    436 aggaaaatggcaaagctgatgggaaggaggagcagaaaaatcaagagaaggcagtagctg 495

Qy    461 atgaagaggatgaagtctgattctgacggaatggatgaagatgattctgatggtgaggatt 520
Db    496 ctacagcttcaaaatcaagtcttggccttgaaaagaaaagcaaggatgactctgatgatt 555

Qy    521 ctgaggaagaagagcct 537
Db    556 ctgatgaggatgagtct 572
```

RESULT 3
US-09-282-305-13
; Sequence 13, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
; APPLICANT: Baldwin, Donald A.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Crane, Virginia C.
; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
; FILE REFERENCE: 5718-44,
; CURRENT APPLICATION NUMBER: US/09/282,305
; CURRENT FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 60/080,563
; PRIOR FILING DATE: 1998-04-03
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (63)..(971)
US-09-282-305-13

Query Match 10.6%; Score 100; DB 4; Length 1191;
Best Local Similarity 54.9%; Pred. No. 3.3e-18;
Matches 268; Conservative 0; Mismatches 205; Indels 15; Gaps 3;

```
Qy     49 atggagttctggggaattgaagttaaatcaggaaagccagttacagtgactcctgaagaa 108
Db     63 atggagttctggggtctcgaggtcaagcctggttccactgttaagtgtgagcctggatat 122

Qy    109 ggcattcttatccacgtttctcaggcatcgcttggagaatgtaaaaacaagaaggggaqag 168
Db    123 ggctttgtgctgcacctttcccaggctgctcttggggaatcg------aagaagagtgat 176

Qy    169 tttgtgcctttacatgtaaaggttgggaaccagaacttggttctgggaactctatcgact 228
Db    177 aatgccttgatgtatgtcaaaattgatgatcagaaacttgccattggaaccctctctgtt 236

Qy    229 gagaacatccctcagcttttctgtgatttggtattcgacaaggagtttgagctttctcac 288
Db    237 gacaagaacccacacattcaatttgatctgattttcgataaagagtttgagctgtcgcac 296

Qy    289 acttggggaaaaggaagtgtttactttgttggataca------aaactcccaacattgag 342
Db    297 acatcaaaaactaccagcgtcttcttcactggctacaaggttgaacagccattcgaggaa 356

Qy    343 ccacaaggctattctgaggaagaagaggaagaagaggaagaagttcctgctgggaatgct 402
Db    357 gatgaaatggatcttgattctgaagatgaagacgaggagctgaatgttccagtagtcaag 416

Qy    403 gccaaggctgtagctaaaccaaaggctaagcctgcagaagtgaagccagctgttgatgat 462
Db    417 gaaaatggcaaagctgatgggaagaaacagaaaagtcaagaaaaggcagttgctgcacct 476
```

```
Qy    463 gaagaggatgagtctgattc---tgacggaatggatgaagatgattctgatggtgaggat 519
          | |        | |||||     |  | | |||||| ||||||||||||||  || ||
Db    477 tcaaaatcaagtccggattccaagaagagcaaggatgacgatgattctgatgaggacgag 536

Qy    520 tctgagga 527
          |||| ||
Db    537 actgatga 544


RESULT   4
US-09-282-305-11
; Sequence 11, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
;  APPLICANT:  Baldwin, Donald A.
;  APPLICANT:  Briggs, Steven P.
;  APPLICANT:  Crane, Virginia C.
;  TITLE OF INVENTION:  Maize Histone Deacetylases And Their Uses
;  FILE REFERENCE:  5718-44,
;  CURRENT APPLICATION NUMBER: US/09/282,305
;  CURRENT FILING DATE:  1999-03-31
;  PRIOR APPLICATION NUMBER:  60/080,563
;  PRIOR FILING DATE:  1998-04-03
;  NUMBER OF SEQ ID NOS:  18
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 11
;    LENGTH: 1283
;    TYPE: DNA
;    ORGANISM: Zea mays
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (79)..(996)
US-09-282-305-11


  Query Match             10.5%;  Score 98.4;  DB 4;  Length 1283;
  Best Local Similarity   54.7%;  Pred. No. 9.5e-18;
  Matches  267;  Conservative    0;  Mismatches  206;  Indels   15;  Gaps    3;

Qy     49 atggagttctggggaattgaagttaaatcaggaaagccagttacagtgactcctgaagaa 108
          ||||||||||||||  | || || ||  | ||     | ||||        |||| | |
Db     79 atggagttctggggtctcgaggtcaagcctggttccactgttaagtgtgagcctggatat 138

Qy    109 ggcattcttatccacgtttctcaggcatcgcttggagaatgtaaaaacaagaagggagag 168
          ||| || |  | ||| |||| |||||  | |||||  ||||        ||||||  | ||
Db    139 ggctttgtgctgcacctttcccaggctgctcttggggaatcg------aagaagagtgat 192

Qy    169 tttgtgcctttacatgtaaaggttgggaaccagaacttggttctgggaactctatcgact 228
            ||       |  |||| ||  |||   | |||||  | |   | |||| || ||   |
Db    193 aatgccttgatgtatgtcaaaattgatgatcagaaacttgccattggaaccctctctgtt 252

Qy    229 gagaacatccctcagcttttctgtgatttggtattcgacaaggagtttgagctttctcac 288
          || || | ||| ||  |    | |||| || | ||||| || ||||||||||| || |||
Db    253 gacaagaacccacacatacaatttgatctgattttcgataaagagtttgagctgtcgcac 312

Qy    289 acttggggaaaaggaagtgtttactttgttggatacaaaact------cccaacattgag 342
          || |   ||     || |||| |||   ||| |||||   |      ||   |   ||
Db    313 acatcaaaaactaccagcgttttcttcactggctacaaggttgagcagccattcgaggaa 372

Qy    343 ccacaaggctattctgaggaagaagaggaagaagaggaagaagttcctgctgggaatgct 402
            ||   ||  |||    ||||| ||||| |||||       |  | | |
Db    373 gatgaaatggatcttgattctgaagatgaagacgaggagctgaatgttccagtagtcaag 432

Qy    403 gccaaggctgtagctaaaccaaaggctaagcctgcagaagtgaagccagctgttgatgat 462
          |  || |    |||| |    |||    ||       |||  ||| ||| || ||    |
Db    433 gaaaatggcaaagctgatgagaagaaacagaaaagtcaagaaaaggcagttgctgcacct 492

Qy    463 gaagaggatgagtctgattc---tgacggaatggatgaagatgattctgatggtgaggat 519
          | |        | |||||     |  | | |||||| || ||||||||||  || ||
Db    493 tcaaaatcaagtccggattccaagaagagcaaggatgacgacgattctgatgaggacgag 552

Qy    520 tctgagga 527
          |||| ||
Db    553 actgatga 560


RESULT   5
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14


  Query Match             8.2%;  Score 77;  DB 1;  Length 7218;
  Best Local Similarity   5.6%;  Pred. No. 1.6e-11;
  Matches   23;  Conservative  238;  Mismatches  148;  Indels    0;  Gaps    0;

Qy    267 caaggagtttgagctttctcacacttggggaaaaggaagtgtttacttttgttggatacaa 326
          |||| ||||  ||   |   | | || || | ::::::: :   :    :  ::: : ::
Db   1466 CAAGTAGTTAAAGAGATAGAAGAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1407

Qy    327 aactcccaacattgagccacaaggctattctgaggaagaagaggaagaagaggaagaagt 386
          ::     :: :  :::  : ::::  :    ::::::::::::::::::::::::::::
Db   1406 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1347

Qy    387 tcctgctgggaatgctgccaaggctgtagctaaaccaaaggctaagcctgcagaagtgaa 446
            :  ::::: :  :  :::  : ::  :::  :::::  :::   : ::::: :::
Db   1346 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1287

Qy    447 gccagctgttgatgatgaagaggatgagtctgattctgacggaatggatgaagatgattc 506
          :  ::  :  :: :: :::::::: :::   ::    :: :::: ::: ::::: ::
Db   1286 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1227
```

```
Qy     507 tgatggtgaggattctgaggaagaagagcctacacctaagaagcctgcatcaagcaagaa 566
           :: :: :::::    ::::::::::::   : :   ::::::   : :  ::: :::::
Db    1226 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1167

Qy     567 gagagctaatgaaactacccctaaagcacctgtgtcagcaaagaaggcgaaaqtagcagt 626
           :::::  :: ::::  :     :::: :   : :  :: :::::::: ::.:: :: ::
Db    1166 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1107

Qy     627 tactcctcagaaaacagatgagaagaagaaaggggggaaaggctgcaaac 675
           :      :::::: ::: :::::::::::::::::::::   |||| |
Db    1106 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRATCGCAAGC 1058


RESULT   6
US-09-433-699-3
; Sequence 3, Application US/09433699B
; Patent No. 6165786
; GENERAL INFORMATION:
;  APPLICANT: C. Frank Bennett
;  APPLICANT: Lex M. Cowsert
;  TITLE OF INVENTION: ANTISENSE MODULATION OF NUCLEOLIN EXPRESSION
;  FILE REFERENCE: RTS-0109
;  CURRENT APPLICATION NUMBER: US/09/433,699B
;  CURRENT FILING DATE: 1999-11-03
;  NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 3
;   LENGTH: 2518
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (112)..(2235)
US-09-433-699-3


  Query Match             6.2%;  Score 57.8;  DB 4;  Length 2518;
  Best Local Similarity   49.5%;  Pred. No. 1.9e-06;
  Matches  149;  Conservative    0;  Mismatches  152;  Indels    0;  Gaps    0;

Qy     347 aaggctattctgaggaagaagaggaagaagaggaagaagttcctgctgggaatgctgcca 406
           |||   |   ||| ||    ||||| || ||||| || |     |  | | |||  |
Db     551 aagaggaggatgatgacagtgaggaggatgaggaggatgacgaggacgaggatgaggatg 610

Qy     407 aggctgtagctaaaccaaaggctaagcctgcagaagtgaagccagctgttgatgatgaag 466
           | | || |  | |||||   || | |   |||| ||      |  |||     || || |
Db     611 aagatgaaattgaaccagcagcgatgaaagcagcagctgctgcccctgcctcaqaggatg 670

Qy     467 aggatgagtctgattctgacggaatggatgaagatgattctgatggtgaggattctgagg 526
           |||| ||    |||   || |     || || || |||   ||||  || |||      |
Db     671 aggacgatgaggatgacgaagatgatgaggatgacgatgacgatgaggaagatgactctg 730

Qy     527 aaqaagagcctacacctaaqaaqcctgcatcaagcaagaagagagctaatgaaactaccc 586
           |||||||  |||     |  |  || ||   | | |||||    || || |    | |
Db     731 aagaagaagctatggagactacaccagccaaaggaaagaaagctgcaaaagttgttcctg 790

Qy     587 ctaaagcacctgtgtcagcaaagaaggcgaaagtagcagttactcctcagaaaacagatg 646
             |||||          ||  || | |   | | || ||    |  | || |    ||||
Db     791 tqaaagccaagaacgtggctgaggatgaagatgaagaagaggatgatgaggacgaggatg 850

Qy     647 a 647
           |
Db     851 a 851


RESULT   7
US-08-574-959A-8
; Sequence 8, Application US/08574959A
; Patent No. 5962224
;  GENERAL INFORMATION:
;    APPLICANT:  Jaekyoon Shin, Insil Joung, Ratna K. Vadlamudi
;    APPLICANT:  and Jack L. Strominger
;    TITLE OF INVENTION:  p62 POLYPEPTIDES, RELATED POLYPEPTIDES
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, Suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109-1875
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Patentin Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/574,959A
;      FILING DATE:  19-DEC-95
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mandragouras, Amy E.
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  DFN-008
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)227-5941
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3211 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  439..3157
US-08-574-959A-8


  Query Match             6.0%;  Score 56.6;  DB 2;  Length 3211;
  Best Local Similarity   57.7%;  Pred. No. 4.4e-06;
  Matches  101;  Conservative    0;  Mismatches   74;  Indels    0;  Gaps    0;

Qy     358 gaggaagaagaggaagaagaggaagaagttcctgctgggaatgctgccaaggctgtagct 417
           ||||||||||  ||||| || |||||||      |  |   | |  |   | |  | ||
Db    2407 GAGGAAGAAGGAGAAGAGGAAGAAGAAGAAGAAGAAGAAGAGGAAGAAGAAGAAGAG 2466

Qy     418 aaaccaaaggctaagcctgcagaagtgaagccagctgttgatgatgaaqaggatgaqtct 477
           ||  | |||   ||   | ||| |  ||  ||  |  || ||||||||||  | | |
Db    2467 GAAGAAGAGGAAGAGGAGGAAGACTTTGAGGAAGAGGAAGAGGATGAAGAGGAATATTTT 2526

Qy     478 gattctgacggaatggatgaagatgattctgatggtgaggattctgaggaagaag 532
           ||    || | |  ||| ||||| || | ||| |  ||   |   || |||||||
Db    2527 GAAGAGGAAGAAGAGGAGGAAGAAGAGTTTGAGGAAGAATTTGAGGAAGAAGAAG 2581


RESULT   8
US-09-357-014-8
; Sequence 8, Application US/09357014
; Patent No. 6291645
;  GENERAL INFORMATION:
;       APPLICANT: Jaekyoon Shin, Insil Joung, Ratna K. Vadlamudi
;                  and Jack L. Strominger
;       TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
;                           AND USES THEREFOR
;       NUMBER OF SEQUENCES: 22
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: LAHIVE & COCKFIELD
;            STREET: 60 State Street, Suite 510
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: USA
;            ZIP: 02109-1875
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
```

```
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;         CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/357,014
;             FILING DATE: 19-Jul-1999
;         PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/574,959
;             FILING DATE: <Unknown>
;         ATTORNEY/AGENT INFORMATION:
;             NAME: Mandragouras, Amy E.
;             REGISTRATION NUMBER: 36,207
;             REFERENCE/DOCKET NUMBER: DFN-008
;         TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617)227-7400
;             TELEFAX: (617)227-5941
;    INFORMATION FOR SEQ ID NO: 8:
;         SEQUENCE CHARACTERISTICS:
;             LENGTH: 3211 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;         MOLECULE TYPE: cDNA
;         FEATURE:
;             NAME/KEY: CDS
;             LOCATION: 439..3157
;         SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-357-014-8


  Query Match             6.0%;  Score 56.6;  DB 4;  Length 3211;
  Best Local Similarity   57.7%;  Pred. No. 4.4e-06;
  Matches  101;  Conservative    0;  Mismatches   74;  Indels    0;  Gaps    0;

Qy     358 gaggaagaagaggaagaagaggaagaagttcctgctgggaatgctgccaaggctgtagct 417
           ||||||||||  ||||| || |||||||      |  |   | |  |   | |  | ||
Db    2407 GAGGAAGAAGGAGAAGAGGAAGAAGAAGAAGAAGAAGAAGGAAGAAGAAGAAGAG 2466

Qy     418 aaaccaaaggctaagcctgcagaagtgaagccagctgttgatgatgaagaggatgagtct 477
             ||  | |||   ||   | |||  |  ||  ||  |  || ||||||||||  | | |
Db    2467 GAAGAAGAGGAAGAGGAGGAAGACTTTGAGGAAGAGGAAGAGGATGAAGAGGAATATTTT 2526

Qy     478 gattctgacggaatggatgaagatgattctgatggtgaggattctgaggaagaag 532
           ||    || | |  ||| ||||| || | ||| |  ||    |   || |||||||
Db    2527 GAAGAGGAAGAAGAGGAGGAAGAAGAGTTTGAGGAAGAATTTGAGGAAGAAGAAG 2581


RESULT  9
US-08-574-959A-6
; Sequence 6, Application US/08574959A
; Patent No. 5962224
;  GENERAL INFORMATION:
;    APPLICANT:  Jaekyoon Shin, Insil Joung, Ratna K. Vadlamudi
;    APPLICANT:  and Jack L. Strominger
;    TITLE OF INVENTION:  p62 POLYPEPTIDES, RELATED POLYPEPTIDES
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, Suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109-1875
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/574,959A
;      FILING DATE:  19-DEC-95
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mandragouras, Amy E.
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  DFN-008
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)227-5941
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3901 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  439..3847
US-08-574-959A-6


  Query Match             6.0%;  Score 56.6;  DB 2;  Length 3901;
  Best Local Similarity   57.7%;  Pred. No. 4.9e-06;
  Matches  101;  Conservative    0;  Mismatches   74;  Indels    0;  Gaps    0;

Qy     358 gaggaagaagaggaagaagaggaagaagttcctgctgggaatgctgccaaggctgtagct 417
           ||||||||||  ||||| || |||||||      |  |   | |  |   | |  | ||
Db    3097 GAGGAAGAAGGAGAAGAGGAAGAAGAAGAAGAAGAAGAAGGAAGAAGAAGAAGAG 3156

Qy     418 aaaccaaaggctaagcctgcagaagtgaagccagctgttgatgatgaagaggatgagtct 477
             ||  | |||   ||   | |||  |  ||  ||  |  || ||||||||||  | | |
Db    3157 GAAGAAGAGGAAGAGGAGGAAGACTTTGAGGAAGAGGAAGAGGATGAAGAGGAATATTTT 3216

Qy     478 gattctgacggaatggatgaagatgattctgatggtgaggattctgaggaagaag 532
           ||    || | |  ||| ||||| || | ||| |  ||    |   || |||||||
Db    3217 GAAGAGGAAGAAGAGGAGGAAGAAGAGTTTGAGGAAGAATTTGAGGAAGAAGAAG 3271


RESULT  10
US-09-357-014-6
; Sequence 6, Application US/09357014
; Patent No. 6291645
;  GENERAL INFORMATION:
;      APPLICANT: Jaekyoon Shin, Insil Joung, Ratna K. Vadlamudi
;                 and Jack L. Strominger
;      TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
;                          AND USES THEREFOR
;      NUMBER OF SEQUENCES: 22
;      CORRESPONDENCE ADDRESS:
;           ADDRESSEE: LAHIVE & COCKFIELD
;           STREET: 60 State Street, Suite 510
;           CITY: Boston
;           STATE: Massachusetts
;           COUNTRY: USA
;           ZIP: 02109-1875
;      COMPUTER READABLE FORM:
;           MEDIUM TYPE: Floppy disk
;           COMPUTER: IBM PC compatible
;           OPERATING SYSTEM: PC-DOS/MS-DOS
;           SOFTWARE: PatentIn Release #1.0, Version #1.25
;      CURRENT APPLICATION DATA:
;           APPLICATION NUMBER: US/09/357,014
;           FILING DATE: 19-Jul-1999
;      PRIOR APPLICATION DATA:
;           APPLICATION NUMBER: 08/574,959
;           FILING DATE: <Unknown>
;      ATTORNEY/AGENT INFORMATION:
;           NAME: Mandragouras, Amy E.
;           REGISTRATION NUMBER: 36,207
;           REFERENCE/DOCKET NUMBER: DFN-008
;      TELECOMMUNICATION INFORMATION:
;           TELEPHONE: (617)227-7400
;           TELEFAX: (617)227-5941
;  INFORMATION FOR SEQ ID NO: 6:
;      SEQUENCE CHARACTERISTICS:
;           LENGTH: 3901 base pairs
```

```
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;          MOLECULE TYPE: cDNA
;          FEATURE:
;             NAME/KEY:  CDS
;             LOCATION:  439..3847
;          SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-357-014-6


  Query Match             6.0%;  Score 56.6;  DB 4;  Length 3901;
  Best Local Similarity   57.7%;  Pred. No. 4.9e-06;
  Matches  101;  Conservative    0;  Mismatches   74;  Indels    0;  Gaps    0;

Qy     358 gaggaagaagaggaagaagaggaagaagttcctgctgggaatgctgccaaggctgtagct 417
           ||||||||||  ||||| || |||||||     |  |   | |  |   | |  | ||
Db    3097 GAGGAAGAAGGAGAAGAGGAAGAAGAAGAAGAAGAAGAAGAGGAAGAAGAAGAAGAG 3156

Qy     418 aaaccaaaggctaagcctgcagaagtgaagccagctgttgatgatgaagaggatgagtct 477
           ||  | |||   ||   | |||  |  ||  ||  |  || ||||||||||||  | | |
Db    3157 GAAGAAGAGGAAGAGGAGGAAGACTTTGAGGAAGAGGAAGAGGATGAAGAGGAATATTTT 3216

Qy     478 gattctgacggaatggatgaagatgattctgatggtgaggattctgaggaagaag 532
           ||    || | |  ||| ||||| || | ||| |  ||   |   || |||||||
Db    3217 GAAGAGGAAGAAGAGGAGGAAGAAGAGTTTGAGGAAGAATTTGAGGAAGAAGAAG 3271


RESULT  11
US-09-538-871-3
; Sequence 3, Application US/09538871
; Patent No. 6358718
; GENERAL INFORMATION:
;  APPLICANT: Olson, Mark
;  APPLICANT: Szebeni, Atilla
;  TITLE OF INVENTION: METHODS FOR STABILIZATION AND
;  TITLE OF INVENTION: RENATURATION OF PROTEINS USING NUCLEOLAR PROTEINS B23
;  FILE REFERENCE: 13121.0002U2
;  CURRENT APPLICATION NUMBER: US/09/538,871
;  CURRENT FILING DATE: 2000-03-30
;  EARLIER APPLICATION NUMBER: 60/126,910
;  EARLIER FILING DATE: 1999-03-30
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 1164
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of artificial sequence:/No. 6358718e -
;   OTHER INFORMATION: synthetic construct
US-09-538-871-3


  Query Match             5.9%;  Score 55.6;  DB 4;  Length 1164;
  Best Local Similarity   50.8%;  Pred. No. 5.2e-06;
  Matches  133;  Conservative    0;  Mismatches  129;  Indels    0;  Gaps    0;

Qy     292 tggggaaaaggaagtgtttactttgttggatacaaaactcccaacattgagccacaaggc 351
           || ||    ||   |||  || |   |||| |  |           | |||  | | |
Db     385 tgtggttctgggcctgtgcacataagtggacagcacctagtagctgtagaggaagatgca 444

Qy     352 tattctgaggaagaagaggaagaagaggaagaagttcctgctgggaatgctgccaaggct 411
           | || || || ||||| || ||||| | | || |     |   ||     |   ||
Db     445 gagtcagaagatgaagatgaggaagatgtaaaactcttaggcatgtctggaaagagatct 504

Qy     412 gtagctaaaccaaaggctaagcctgcagaagtgaagccagctgttgatgatgaagaggat 471
           |   |   |         ||     || |    ||   |    ||||||| || || |||
Db     505 gctcccggaggtggtaacaaagtcccacagaaaaaagtaaaacttgatgaagatgatgat 564

Qy     472 gagtctgattctgacggaatggatgaagatgattctgatggtgaggattctgaggaagaa 531
           |||  ||||   || |    ||||||| || |||  ||||| ||| |||| ||| ||||||
Db     565 gaggatgatgaagatgatgaggatgatgaagatgatgatgatgatgattttgatgaggag 624

Qy     532 gagcctacacctaagaagcctg 553
           ||  ||  |   |||   || |
Db     625 gaaactgaagaaaaggttccag 646


RESULT  12
US-09-538-871-1
; Sequence 1, Application US/09538871
; Patent No. 6358718
; GENERAL INFORMATION:
;  APPLICANT: Olson, Mark
;  APPLICANT: Szebeni, Atilla
;  TITLE OF INVENTION: METHODS FOR STABILIZATION AND
;  TITLE OF INVENTION: RENATURATION OF PROTEINS USING NUCLEOLAR PROTEINS B23
;  FILE REFERENCE: 13121.0002U2
;  CURRENT APPLICATION NUMBER: US/09/538,871
;  CURRENT FILING DATE: 2000-03-30
;  EARLIER APPLICATION NUMBER: 60/126,910
;  EARLIER FILING DATE: 1999-03-30
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 1232
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of artificial sequence:/No. 6358718e -
;   OTHER INFORMATION: synthetic construct
US-09-538-871-1


  Query Match             5.9%;  Score 55.6;  DB 4;  Length 1232;
  Best Local Similarity   50.8%;  Pred. No. 5.3e-06;
  Matches  133;  Conservative    0;  Mismatches  129;  Indels    0;  Gaps    0;

Qy     292 tggggaaaaggaagtgtttactttgttggatacaaaactcccaacattgagccacaaggc 351
           || ||    ||   |||  || |   |||| |  |           | |||  | | |
Db     356 tgtggttctgggcctgtgcacataagtggacagcacctagtagctgtagaggaagatgca 415

Qy     352 tattctgaggaagaagaggaagaagaggaagaagttcctgctgggaatgctgccaaggct 411
           | || || || ||||| || ||||| | | || |     |   ||     |   ||
Db     416 gagtcagaagatgaagatgaggaagatgtaaaactcttaggcatgtctggaaagagatct 475

Qy     412 gtagctaaaccaaaggctaagcctgcagaagtgaagccagctgttgatgatgaagaggat 471
           |   |   |         ||     || |    ||   |    ||||||| || || |||
Db     476 gctcccggaggtggtaacaaagtcccacagaaaaaagtaaaacttgatgaagatgatgat 535

Qy     472 gagtctgattctgacggaatggatgaagatgattctgatggtgaggattctgaggaagaa 531
           |||  ||||   || |    ||||||| || |||  ||||| ||| |||| ||| ||||||
Db     536 gaggatgatgaagatgatgaggatgatgaagatgatgatgatgatgattttgatgaagag 595

Qy     532 gagcctacacctaagaagcctg 553
           ||  ||  |   |||   || |
Db     596 gaaactgaagaaaaggttccag 617


RESULT  13
US-08-728-323A-1
; Sequence 1, Application US/08728323A
; Patent No. 5948676
;  GENERAL INFORMATION:
;    APPLICANT:  Chang, Yuan
;    APPLICANT:  Bohenzky, Roy A.
;    APPLICANT:  Russo, James J.
;    APPLICANT:  Edelman, Isidore S.
;    APPLICANT:  Moore, Patrick S.
;    TITLE OF INVENTION:  Immediate Early Protein From Kaposi's
;    TITLE OF INVENTION:  Sarcoma-Associated Herpesvirus, DNA
;    TITLE OF INVENTION:  Encoding Same And Uses Thereof
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
```

```
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/728,323A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  0575/52268/JPW/MSC/SKS
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-278-0400
;      TELEFAX:  212-391-0525
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3489 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..3489
US-08-728-323A-1


  Query Match             5.6%;  Score 52.2;  DB 2;  Length 3489;
  Best Local Similarity   55.9%;  Pred. No. 7.5e-05;
  Matches   99;  Conservative    0;  Mismatches   78;  Indels    0;  Gaps    0;

Qy     357 tgaggaagaagaggaagaagaggaagaagttcctgctgggaatgctgccaaggctgtagc 416
           |||||| || ||||| || || || || | |   | | | |||  |   ||| || |
Db    1014 TGAGGAGGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGA 1073

Qy     417 taaaccaaaggctaagcctgcagaagtgaagccagctgttgatgatgaagaggatgagtc 476
             |    | |   ||  ||  || | | |    | |  ||||| || ||||||||
Db    1074 GGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGA 1133

Qy     477 tgattctgacggaatggatgaagatgattctgatggtgaggattctgaggaagaaga 533
            ||   |||||    |||||| || ||    || |  |||||    ||||| || ||
Db    1134 GGAGGATGACGAGGAGGATGACGAGGAGGAGGACGAGGAGGAGGACGAGGAGGAGGA 1190


RESULT  14
US-08-770-379-20/c
; Sequence 20, Application US/08770379
; Patent No. 5849564
;  GENERAL INFORMATION:
;    APPLICANT:  Chang, Yuan
;    APPLICANT:  Bohenzky, Roy A.
;    APPLICANT:  Russo, James J.
;    APPLICANT:  Edelman, Isidore S.
;    APPLICANT:  Moore, Patrick S.
;    TITLE OF INVENTION:  POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED
;    TITLE OF INVENTION:  HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/770,379
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  52342
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 278-0400
;      TELEFAX:  (212) 391-0525
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  32207 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-770-379-20


  Query Match             5.6%;  Score 52.2;  DB 2;  Length 32207;
  Best Local Similarity   55.9%;  Pred. No. 0.00021;
  Matches   99;  Conservative    0;  Mismatches   78;  Indels    0;  Gaps    0;

Qy     357 tgaggaagaagaggaagaagaggaagaagttcctgctgggaatgctgccaaggctgtagc 416
           |||||| || ||||| || || || || | |   | | | |||  |   ||| || |
Db   20983 TGAGGAGGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGA 20924

Qy     417 taaaccaaaggctaagcctgcagaagtgaagccagctgttgatgatgaagaggatgagtc 476
             |    | |   ||  ||  || | | |    | |  ||||| || ||||||||
Db   20923 GGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGA 20864

Qy     477 tgattctgacggaatggatgaagatgattctgatggtgaggattctgaggaagaaga 533
            ||   |||||    |||||| || ||    || |  |||||    ||||| || ||
Db   20863 GGAGGATGACGAGGAGGATGACGAGGAGGAGGACGAGGAGGAGGACGAGGAGGAGGA 20807


RESULT  15
US-08-757-669A-20/c
; Sequence 20, Application US/08757669A
; Patent No. 6183751
;  GENERAL INFORMATION:
;    APPLICANT:  Chang, Yuan
;    APPLICANT:  Bohenzky, Roy A.
;    APPLICANT:  Russo, James J.
;    APPLICANT:  Edelman, Isidore S.
;    APPLICANT:  Moore, Patrick S.
;    TITLE OF INVENTION:  UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
;    TITLE OF INVENTION:  SEQUENCES AND USES THEREOF
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/757,669A
;      FILING DATE:
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
```

```
;      REFERENCE/DOCKET NUMBER:  45185-F
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 278-0400
;      TELEFAX:  (212) 391-0525
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  32207 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-757-669A-20


  Query Match             5.6%;  Score 52.2;  DB 4;  Length 32207;
  Best Local Similarity   55.9%;  Pred. No. 0.00021;
  Matches   99;  Conservative    0;  Mismatches   78;  Indels    0;  Gaps    0;

Qy     357 tgaggaagaagaggaagaagaggaagaagttcctgctgggaatgctgccaaggctgtagc 416
           |||||| || ||||| || || || || | |   |  | | |||  |   ||| ||  |
Db   20983 TGAGGAGGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGA 20924

Qy     417 taaaccaaaggctaagcctgcagaagtgaagccagctgttgatgatgaagaggatgagtc 476
             |     | |   ||  ||  || | | |    |  |  ||||| || ||||||||
Db   20923 GGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGA 20864

Qy     477 tgattctgacggaatggatgaagatgattctgatggtgaggattctgaggaagaaga 533
           ||   |||||    |||||| || ||    || |  |||||    ||||| || ||
Db   20863 GGAGGATGACGAGGAGGATGACGAGGAGGAGGACGAGGAGGAGGACGAGGAGGAGGA 20807


Search completed: May  2, 2002, 20:15:18
Job time: 15105 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 15:42:36 ; Search time 3265.32 Seconds
(without alignments)
3881.286 Million cell updates/sec

Title: US-09-645-337-5
Perfect score: 939
Sequence: 1 cacgcgtccgtaaaaatcct.....aaaaaaaaaagggcggccgc 939

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	636	67.7	636	9	AV544477	AV544477 AV544477
	2	451.2	48.1	534	9	AI995198	AI995198 701502609
	3	367.6	39.1	434	10	N37851	N37851 19078 Lambd
	4	156.2	16.6	422	10	BE523968	BE523968 M44A9STM
	5	156.2	16.6	728	10	BE844690	BE844690 AD01A07T7
	6	149.8	16.0	504	10	T45874	T45874 9137 Lambda
	7	146.8	15.6	543	10	BF279194	BF279194 GA__Eb003
	8	145.8	15.5	378	9	AV551556	AV551556 AV551556
	9	141.6	15.1	759	10	BG131117	BG131117 EST464009
	10	141.6	15.1	875	10	BF275467	BF275467 GA__Eb002
	11	141	15.0	327	9	AV549937	AV549937 AV549937
	12	140.2	14.9	511	10	BI130765	BI130765 G110P49Y
	13	137.8	14.7	865	10	BG445639	BG445639 GA__Ea002
	14	136.2	14.5	517	10	BF278973	BF278973 GA__Eb003
	15	136	14.5	715	10	BF270691	BF270691 GA__Eb000
	16	135.6	14.4	382	10	C99809	C99809 C99809 YAC
	17	131.2	14.0	560	10	BI786647	BI786647 sai51f09.
	18	130.4	13.9	537	10	BM177387	BM177387 saj80q01.
	19	129.8	13.8	539	9	AW906751	AW906751 EST342874
	20	129.2	13.8	595	9	AW929159	AW929159 EST337947
	21	128.6	13.7	479	10	BI924066	BI924066 EST543955
	22	128.6	13.7	504	10	BG134445	BG134445 EST467337
	23	128.6	13.7	537	9	AW615938	AW615938 EST325304
	24	128.6	13.7	542	9	AW929950	AW929950 EST354220
	25	128.6	13.7	552	10	BG134979	BG134979 EST467871
	26	128.6	13.7	555	10	BE449829	BE449829 EST361267
	27	128.6	13.7	589	10	BI934377	BI934377 EST554266
	28	128.6	13.7	590	10	BI930352	BI930352 EST550241
	29	128.6	13.7	611	10	BI924005	BI924005 EST543894
	30	128.6	13.7	668	10	BI927751	BI927751 EST547640
	31	128	13.6	674	10	BG595734	BG595734 EST494412
	32	127.6	13.6	435	9	AW429116	AW429116 EST306572
	33	127.6	13.6	579	10	BE343018	BE343018 EST395862
	34	127.6	13.6	591	10	BE343072	BE343072 EST395916
	35	127.6	13.6	661	10	BF460278	BF460278 073G03 Ma
	36	127	13.5	453	10	BF187099	BF187099 EST443386
	37	127	13.5	467	9	AI776222	AI776222 EST257322
	38	127	13.5	467	10	BE342419	BE342419 EST395263
	39	127	13.5	473	10	BE922565	BE922565 EST426334
	40	127	13.5	548	10	BG594032	BG594032 EST492710
	41	127	13.5	578	10	BI433232	BI433232 EST535993
	42	126.6	13.5	387	9	AW039158	AW039158 EST281393
	43	125.8	13.4	686	10	BG594317	BG594317 EST492995
	44	125.6	13.4	581	10	BG889339	BG889339 EST515190
	45	125.6	13.4	635	10	BG597002	BG597002 EST495680

ALIGNMENTS

RESULT 1
AV544477/c
LOCUS AV544477 636 bp mRNA linear EST 06-SEP-2000
DEFINITION AV544477 Arabidopsis thaliana roots Columbia Arabidopsis thaliana cDNA clone RZ45c06F 3', mRNA sequence.
ACCESSION AV544477
VERSION AV544477.1 GI:8715891
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 636)
AUTHORS Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
JOURNAL DNA Res. 7, 175-180 (2000)
MEDLINE 20363093
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES Location/Qualifiers
source 1. .636
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="RZ45c06F"
/clone_lib="Arabidopsis thaliana roots Columbia"
/tissue_type="roots"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 147 a 167 c 121 g 201 t
ORIGIN

Query Match 67.7%; Score 636; DB 9; Length 636;

```
  Best Local Similarity   100.0%;  Pred. No. 3.7e-112;
  Matches  636;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     220 ctatcgactgagaacatccctcagcttttctgtgatttggtattcgacaaggagtttgag 279
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     636 CTATCGACTGAGAACATCCCTCAGCTTTTCTGTGATTTGGTATTCGACAAGGAGTTTGAG 577

Qy     280 ctttctcacacttqqqgaaaaqgaaqtgtttacttttgttggatacaaaactcccaacatt 339
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     576 CTTTCTCACACTTGGGGAAAAGGAAGTGTTTACTTTGTTGGATACAAAACTCCCAACATT 517

Qy     340 gagccacaaggctattctgaggaagaagaggaagaagaggaagaagttcctgctgggaat 399
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     516 GAGCCACAAGGCTATTCTGAGGAAGAAGAGGAAGAAGAGGAAGAAGTTCCTGCTGGGAAT 457

Qy     400 gctgccaaggctgtagctaaaccaaaggctaagcctgcagaagtgaagccagctgttgat 459
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     456 GCTGCCAAGGCTGTAGCTAAACCAAAGGCTAAGCCTGCAGAAGTGAAGCCAGCTGTTGAT 397

Qy     460 gatgaaagaggatgagtctgattctgacggaatggatgaagatgattctgatggtgaggat 519
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     396 GATGAAGAGGATGAGTCTGATTCTGACGGAATGGATGAAGATGATTCTGATGGTGAGGAT 337

Qy     520 tctgaggaagaagaggcctacacctaagaagcctgcatcaagcaagaagagagctaatgaa 579
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     336 TCTGAGGAAGAAGAGCCTACACCTAAGAAGCCTGCATCAAGCAAGAAGAGAGCTAATGAA 277

Qy     580 actacccctaaagcacctgtgtcagcaaagaaggcgaaagtagcagttactcctcagaaa 639
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     276 ACTACCCCTAAAGCACCTGTGTCAGCAAAGAAGGCGAAAGTAGCAGTTACTCCTCAGAAA 217

Qy     640 acagatgaqaagaagaaaqqqqgaaaggctgcaaaccagagcccaaagtcggccagtcaa 699
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     216 ACAGATGAGAAGAAGAAAGGGGGAAAGGCTGCAAACCAGAGCCCAAAGTCGGCCAGTCAA 157

Qy     700 gtctcatgtggttcatgcaagaagactttcaactcagggaatgcacttgagtctcacaac 759
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     156 GTCTCATGTGGTTCATGCAAGAAGACTTTCAACTCAGGGAATGCACTTGAGTCTCACAAC 97

Qy     760 aaggccaagcacgctgctgccaagtgaagtggtttcttattagagcttgtgatttctatg 819
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      96 AAGGCCAAGCACGCTGCTGCCAAGTGAAGTGGTTTCTTATTAGAGCTTGTGATTTCTATG 37

Qy     820 gaattttgcctgtagtctttatgaaaccttcggatt 855
           |||||||||||||||||||||||||||||||||||||
Db      36 GAATTTTGCCTGTAGTCTTTATGAAACCTTCGGATT 1


RESULT  2
AI995198
LOCUS       AI995198                534 bp    mRNA    linear   EST 08-SEP-1999
DEFINITION  701502609 A. thaliana, Ohio State clone set Arabidopsis thaliana
            cDNA clone 701502609, mRNA sequence.
ACCESSION   AI995198
VERSION     AI995198.1  GI:5842103
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 534)
  AUTHORS   Chen,J., Momiyama,M., Chan,E., Mooney,M., Carroon,B., Gilliland,D.,
            Wang,X., Hillman,J., Guegler,K., Kim,C., Doyle,M., Brzoska,P.,
            Gorgone,G., Burns,D., Griffin,J., Mouanoutoua,M., Nguyen,D., Tan,R.
            , Rose,M., Warren,B., Ton,B., Kastury,K., Borillo,C., Carpio,T.,
            Policky,J., Suzuki,G., Argentine,C., Shah,S., Nobriga,A., Murry,L.,
            Turner,C., Krikorian,S., Elder,L. and Hanson,D.
  TITLE     Arabidopsis thaliana Gene Expression MicroArray
  JOURNAL   Unpublished (1999)
COMMENT     Contact: David Smoller, Ph.D.
            Genome Systems, Inc., a wholly owned subsidiary of Incyte
            Pharmaceuticals, Inc.
            4633 World Parkway Circle, St. Louis, MO 63134, USA
            Tel: 877-577-2733
            Fax: 314-427-3324
            Email: service@genomesystems.com.
FEATURES             Location/Qualifiers
     source          1. .534
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="701502609"
                     /clone_lib="A. thaliana, Ohio State clone set"
                     /note="cDNA library was made from selected clones from the
                     Arabidopsis thaliana Ohio  State clone set."
BASE COUNT      155 a     95 c    143 g    141 t
ORIGIN


  Query Match             48.1%;  Score 451.2;  DB 9;  Length 534;
  Best Local Similarity   98.1%;  Pred. No. 9.8e-77;
  Matches  530;  Conservative    0;  Mismatches    3;  Indels    7;  Gaps    7;

Qy      11 taaaaatcctctcttttttctcaaccttgattcttagccatggagttctggggaattgaag 70
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 TAAAAATCCTCTCTTTTTCTCAACCTTGATTCTTAGCCATGGAGTTCTGGGGAATTGAAG 60

Qy      71 ttaaatcaggaaagccagttacagtgactcctgaagaaggcattcttatccacgtttctc 130
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TTAAATCAGGAAAGCCAGTTACAGTGACTCCTGAAGAAGGCATTCTTATCCACGTTTCTC 120

Qy     131 aggcatcgcttggagaatgtaaaaacaagaagggagagtttgtgccttacatgtaaagg 190
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AGGCATCGCTTGGAGAATGTAAAAACAAGAAGGGAGAGTTTGTGCCTTTACATGTAAAGG 180

Qy     191 ttgggaaccagaacttggttctgggaactctatcgactgagaacatccctcagcttttct 250
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TTGGG-ACCAGAACTTGGTTCTGGGAACTCTATCGACTGAGAACATCCCTCAGCTTTTCT 239

Qy     251 gtgatttggtattcgacaaggagtttgagctttctcacacttggggaaaaggaagtgttt 310
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     240 GTGATTTGGTATTCGACAAGGAGTTTGAGCTTTCTCACACTTGGGGAAAAGGAAGTGTTT 299

Qy     311 actttgttggatacaaaactcccaacattgagccacaaggctattctgaggaagaagagg 370
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     300 ACTTTGTTGGATACAAAACTCCCAACATTGAGCCACAAGGCTATTCTGAGGAAGAAGAGG 359

Qy     371 aagaagaggaagaagttcctgctgggaatgctgccaaggctgtagctaaaccaaaggcta 430
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     360 AAGAAGAGGAAGAAGTTCCTGCTGGGAATGCTGCCAAGGCTGTAGCTAAACCAAAGGCTA 419

Qy     431 agcctgcagaagtgaagccagctgttgatgatgaaqaggatgagtctgattctgacggaa 490
           || |||||||||| |||||||||||||||||| ||||||||||||| ||||||||||||
Db     420 AG-CTGCAGAAGTG-AGCCAGCTGTTGATGATG-AGAGGATGAGTCTG-TTCTGACGGAA 475

Qy     491 tggatgaagatgattctgatggtgaggattctgaggaagaagagcctacacctaagaagc 550
           ||||||| | ||||||||||||||||||||||||| ||| ||| ||||||||||||||||
Db     476 TGGATGACG-TGATTCTGATGGTGAGGATTCTGAGGGAGACGAG-CTACACCTAAGAAGC 533


RESULT  3
N37851
LOCUS       N37851                  434 bp    mRNA    linear   EST 05-JAN-1998
DEFINITION  19078 Lambda-PRL2 Arabidopsis thaliana cDNA clone 211N2T7, mRNA
            sequence.
ACCESSION   N37851
VERSION     N37851.1  GI:1158993
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 434)
  AUTHORS   Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh
            ,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M., Retzel
```

```
          ,E. and Somerville,C.
  TITLE     Genes galore: a summary of methods for accessing results from
            large-scale partial sequencing of anonymous Arabidopsis cDNA clones
  JOURNAL   Plant Physiol. 106, 1241-1255 (1994)
  MEDLINE   95148729
COMMENT     Contact: Thomas Newman
            MSU-DOE Plant Research Laboratory
            Michigan State University
            MSU-DOE-PRL, Michigan State University,Plant Biology Bldg.,E.
            Lansing,Mi
            Tel: 517-353-0854
            Fax: 517-353-9168
            Email: 22313tcn@ibm.cl.msu.edu
            Seq primer: T7 dye primer.
FEATURES             Location/Qualifiers
     source          1. .434
                     /organism="Arabidopsis thaliana"
                     /strain="var columbia"
                     /db_xref="taxon:3702"
                     /clone="211N2T7"
                     /clone_lib="Lambda-PRL2"
                     /note="Vector: lambda Zip-Lox; Site_1: Sal; Site_2: Not;
                     Lambda PRL2 is a cDNA library derived from equal
                     quantities of 4 pools of mRNA.  The mRNA sources were 1) 7
                     day germinated etiolated seedlings; 2) tissue culture
                     grown roots; 3) staged plants half with 24 hour light
                     cycle, half on 16 hr light, 8 hour dark- rosettes;  4)
                     same plants as 3 but aerial tissue (stems, flowers and
                     siliques.  The vector is BRL's lambda Zip-Lox.  The cDNA
                     inserts were directionally cloned with Sal-Not arms using
                     oligo dT primed cDNA.  "
BASE COUNT      117 a     78 c    106 g    116 t     17 others
ORIGIN


  Query Match             39.1%;  Score 367.6;  DB 10;  Length 434;
  Best Local Similarity   92.6%;  Pred. No. 1.1e-60;
  Matches  399;  Conservative    0;  Mismatches   29;  Indels    3;  Gaps    2;

Qy      11 taaaaatcctctcttttttctcaaccttgattcttagccatggagttctggggaattgaag 70
           ||||| |||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       1 TAAAANTCCTCTCTTTTTNTCAACCTTGATTCTTAGCCATGGAGTTCTGGGGAATTGAAG 60

Qy      71 ttaaatcaggaaagccagttacagtgactcctgaagaaggcattcttatccacgtttctc 130
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TTAAATCAGGAAAGCCAGTTACAGTGACTCCTGAAGAAGGCATTCTTATCCACGTTTCTN 120

Qy     131 aggcatcgcttggagaatgtaaaaacaagaagggagagtttgtgcctttacatgtaaagg 190
           |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     121 AGGCATCGCTTGGAGAATGTAAAAACAAGAAGGGAGAGTTTGTNCCTTTACATGTAAAGG 180

Qy     191 ttgggaaccagaacttggttctgggaactctatcgactgagaacatccctcagctttct 250
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||  |
Db     181 TTGGGAACCAGAACTTGGTTCTGGGAACTCTATCGACTGAGAACATCCCTCAGCTTTNNT 240

Qy     251 gtgatttggtattcgacaaggagtttgagctttctcacacttggggaaaaggaagtgtttt 310
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db     241 GTGATTTGGTATTCGACAAGGAGTTTGAGCTTTCTCACACTTGGGGAAAAGGAAGTNTTT 300

Qy     311 actttgttggatacaaaactcccaacattgagccacaaggctattctgaggaagaag--a 368
           ||||||| || |||||||||||||||||||||||||||||||||||||||||||||| ||
Db     301 ACTTTGTNGGGTACAAAACTCCCAACATTGAGCCACAAGGCTATTCTGAGGAAGGAGGNG 360

Qy     369 gqaaqaaqaggaag-aagttcctgctgggaatgctgccaaggctgtagctaaaccaaagg 427
           ||| | |||| || |||||||| ||||| || |||||||| || | |   |||| || |
Db     361 GGAGNAGGAGGNAGNAAGTTCCTNCTGGGGATTCTGCCAAGNCTTTNGGCNAACCCAANG 420

Qy     428 ctaagcctgca 438
           | ||||| ||
Db     421 GTTAGCCTTCA 431


RESULT   4
```

```
BE523968
LOCUS       BE523968                422 bp    mRNA    linear   EST 19-MAR-2001
DEFINITION  M44A9STM Arabidopsis developing seed Arabidopsis thaliana cDNA
            clone M44A9 5', mRNA sequence.
ACCESSION   BE523968
VERSION     BE523968.1  GI:9781946
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 422)
  AUTHORS   White,J.A., Todd,J., Newman,T., Focks,N., Girke,T., Martinez de
            Ilarduya,O., Jaworski,J.G., Ohlrogge,J. and Benning,C.
  TITLE     A new set of Arabidopsis expressed sequence tags from developing
            seeds. The metabolic pathway from carbohydrates to seed oil
  JOURNAL   Plant Physiol. 124 (4), 1582-1594 (2000)
  MEDLINE   20567808
COMMENT     Contact: Benning, C
            Dept. of Biochemistry & Molecular Biology
            Michigan State University
            224 Biochemistry, Michigan State University, East Lansing, MI 48824
            , USA
            Tel: 517 355 1609
            Fax: 517 353 9334
            Email: benning@msu.edu
            Michigan State University DNA Sequencing Facility Arabidopsis
            Biological Resource Center, The Ohio State University, 309 Botany &
            Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210 USA, FAX:
            6142920603 TEL: 6142929371.
FEATURES             Location/Qualifiers
     source          1. .422
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /clone="M44A9"
                     /clone_lib="Arabidopsis developing seed"
                     /tissue_type="seed"
                     /dev_stage="5-13 days after flowering"
                     /lab_host="E.coli"
                     /note="Organ: Developing seed; Vector: pBluescript SK-;
                     Site_1: EcoRI; Site_2: XhoII"
BASE COUNT      109 a     99 c     92 g    122 t
ORIGIN


  Query Match             16.6%;  Score 156.2;  DB 10;  Length 422;
  Best Local Similarity   64.1%;  Pred. No. 3.3e-20;
  Matches  252;  Conservative    0;  Mismatches  138;  Indels    3;  Gaps    1;

Qy      25 ttttctcaaccttgattcttagccatggagttctggggaattgaagttaaatcaggaaag 84
           || ||||  ||| | |    | | ||||||||||||||| |||  || | | ||  |||
Db      31 TTCTCTCTTCCTCGTTCAACAACAATGGAGTTCTGGGGAGTTGCGGTGACACCAAAAAAC 90

Qy      85 ccagttacagtgactcctgaagaaggcattcttatccacgtttctcaggcatcgcttgga 144
             |   ||  |||||||||||||||| ||  ||| ||||| |||||||||| || |||
Db      91 GCTACTAAGGTGACTCCTGAAGAAGACAGCCTTGTCCACATTTCTCAGGCTTCACTT--- 147

Qy     145 gaatgtaaaaacaagaagggagagtttgtgcctttacatgtaaaggttgggaaccagaac 204
           || || | |   ||    ||||| | ||||  |||   ||| |  |||||       ||
Db     148 GACTGCACAGTGAAATCTGGAGAATCTGTGGTTTTGAGTGTGACTGTTGGTGGGGCTAAA 207

Qy     205 ttggttctgggaactctatcgactgagaacatccctcagctttctgtgatttggtattc 264
           | ||| | ||||| || ||    || ||  |||||||| ||  || |||||||| ||
Db     208 CTTGTTATTGGAACACTTTCACAAGACAAGTTCCCTCAGATTAGCTTTGATTTGGTTTTT 267

Qy     265 gacaaggagtttgagctttctcacacttggggaaaaggaagtgtttactttgttggatac 324
           || || ||||||||||||||| ||||   |    |||| || |||| | ||  |||| |||
Db     268 GATAAAGAGTTTGAGCTTTCACACAGCGGTACCAAAGCAAATGTTCATTTCATTGGCTAC 327

Qy     325 aaaactcccaacattgagccacaaggctattctgaggaagaagaggaagaagaggaagaa 384
           ||| | |||||||| ||||   | | ||   ||      || || || || |    |||
```

```
Db     328 AAATCCCCCAACATCGAGCAGGATGACTTCACTAGTTCGGATGATGAGGATGTTCCTGAA 387

Qy     385 gttcctgctgggaatgctgccaaggctgtagct 417
           | |  | |||    |||  | |  |||||  ||
Db     388 GCTGTTCCTGCTCCTGCCCCTACTGCTGTTACT 420


RESULT   5
BE844690
LOCUS       BE844690                728 bp    mRNA    linear   EST 17-MAY-2001
DEFINITION  AD01A07T7 AD A. thaliana (Col-0 gl1) library enriched for
            salt-induced transcripts; 10-14 day seedlings; 4h 160mM NaCl stress
            Arabidopsis thaliana cDNA clone AD01A07 similar to (AF044914)
            putative histone deacetylase, mRNA sequence.
ACCESSION   BE844690
VERSION     BE844690.1  GI:10277068
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 728)
  AUTHORS   Gong,Z., Koiwa,H., Cushman,M.A., Ray,A., Bufford,D., Kore-Eda,S.,
            Matsumoto,T.K., Zhu,J., Cushman,J.C., Bressan,R.A. and Hasegawa
            ,P.M.
  TITLE     Genes that are uniquely stress regulated in salt overly sensitive
            (sos) mutants
  JOURNAL   Plant Physiol. 126 (1), 363-375 (2001)
  MEDLINE   21249177
COMMENT     Contact: Cushman JC
            Department of Biochemistry
            University of Nevada
            MS200, Reno, NV 89557-0014, USA
            Tel: 775-784-1918
            Fax: 775-784-1650
            Email: jcushman@unr.edu
            Plate: AD01  row: A  column: 7
            Seq primer: T7
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .728
                     /organism="Arabidopsis thaliana"
                     /strain="ecotype Columbia gl1"
                     /db_xref="taxon:3702"
                     /clone="AD01A07"
                     /clone_lib="AD A. thaliana (Col-0 gl1) library enriched
                     for salt-induced transcripts; 10-14 day seedlings; 4h
                     160mM NaCl stress"
                     /tissue_type="mixed leaf and root"
                     /dev_stage="10-14 day old"
                     /lab_host="Escherichia coli, TOP10F'"
                     /note="Organ: whole seedlings; Vector: pT-Adv (CLONTECH);
                     Site_1: EcoRI; Site_2: EcoRI; AD Arabidopsis thaliana
                     library enriched for salt-induced transcripts  from
                     wildtype (Col-0 gl1) 10-14 day-old seedlings treated with
                     160 mM  NaCl for 4 hours."
BASE COUNT      219 a    151 c    179 g    179 t
ORIGIN


  Query Match           16.6%;  Score 156.2;  DB 10;  Length 728;
  Best Local Similarity   64.1%;  Pred. No. 2.8e-20;
  Matches  252;  Conservative    0;  Mismatches  138;  Indels    3;  Gaps    1;

Qy      25 ttttctcaaccttgattcttagccatggagttctggggaattgaagttaaatcaggaaag 84
           || ||||  ||| | |    | | ||||||||||||||| |||  || | | ||  |||
Db      29 TTCTCTCTTCCTCGTTCAACAACAATGGAGTTCTGGGGAGTTGCGGTGACACCAAAAAAC 88

Qy      85 ccagttacagtgactcctgaagaaggcattcttatccacgtttctcaggcatcgcttgga 144
             |   ||  |||||||||||||||| ||  ||| ||||| |||||||||| || |||
Db      89 GCTACTAAGGTGACTCCTGAAGAAGACAGCCTTGTCCACATTTCTCAGGCTTCACTT--- 145

Qy     145 gaatgtaaaaacaagaagggagagtttgtgccttacatgtaaaggttgggaaccagaac 204
           || || | |   ||    ||||| | ||||  |||   ||| |  |||||        ||
Db     146 GACTGCACAGTGAAATCTGGAGAATCTGTGGTTTTGAGTGTGACTGTTGGTGGGGCTAAA 205

Qy     205 ttggttctgggaactctatcgactgagaacatccctcagcttttctgtgatttggtattc 264
             | ||| | ||||| || ||    || ||  |||||||| ||  || |||||||| ||
Db     206 CTTGTTATTGGAACACTTTCACAAGACAAGTTCCCTCAGATTAGCTTTGATTTGGTTTTT 265

Qy     265 gacaaggagtttgagctttctcacacttggggaaaaggaagtgtttacttttgttggatac 324
           || || |||||||||||||| ||||   |   |||| || |||| | ||  |||| |||
Db     266 GATAAAGAGTTTGAGCTTTCACACAGCGGTACCAAAGCAAATGTTCATTTCATTGGCTAC 325

Qy     325 aaaactcccaacattgagccacaaggctattctgaggaaagaagaggaagaagaggaagaa 384
           ||| | |||||||| |||| | | ||  ||      || || || || |     |||
Db     326 AAATCCCCCAACATCGAGCAGGATGACTTCACTAGTTCGGATGATGAGGATGTTCCTGAA 385

Qy     385 gttcctgctgggaatgctgccaaggctgtagct 417
           | |  | |||    |||  | |  |||||  ||
Db     386 GCTGTTCCTGCTCCTGCCCCTACTGCTGTTACT 418


RESULT   6
T45874
LOCUS       T45874                  504 bp    mRNA    linear   EST 09-JAN-1998
DEFINITION  9137 Lambda-PRL2 Arabidopsis thaliana cDNA clone 133F15T7, mRNA
            sequence.
ACCESSION   T45874
VERSION     T45874.1  GI:934127
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 504)
  AUTHORS   Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh
            ,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M., Retzel
            ,E. and Somerville,C.
  TITLE     Genes galore: a summary of methods for accessing results from
            large-scale partial sequencing of anonymous Arabidopsis cDNA clones
  JOURNAL   Plant Physiol. 106, 1241-1255 (1994)
  MEDLINE   95148729
COMMENT     On Apr 14, 1993 this sequence version replaced gi:638462.
            Contact: Thomas Newman
            MSU-DOE Plant Research Laboratory
            Michigan State University
            MSU-DOE-PRL, Michigan State University,Plant Biology Bldg.,E.
            Lansing,Mi
            Tel: 517-353-0854
            Fax: 517-353-9168
            Email: 22313tcn@ibm.cl.msu.edu
            Seq primer: T7 dye primer.
FEATURES             Location/Qualifiers
     source          1. .504
                     /organism="Arabidopsis thaliana"
                     /strain="var columbia"
                     /db_xref="taxon:3702"
                     /clone="133F15T7"
                     /clone_lib="Lambda-PRL2"
                     /note="Vector: lambda Zip-Lox; Site_1: Sal; Site_2: Not;
                     Lambda PRL2 is a cDNA library derived from equal
                     quantities of 4 pools of mRNA.  The mRNA sources were 1) 7
                     day germinated etiolated seedlings; 2) tissue culture
                     grown roots; 3) staged plants half with 24 hour light
                     cycle, half on 16 hr light, 8 hour dark- rosettes;  4)
                     same plants as 3 but aerial tissue (stems, flowers and
                     siliques.  The vector is BRL's lambda Zip-Lox.  The cDNA
                     inserts were directionally cloned with Sal-Not arms using
                     oligo dT primed cDNA.  "
BASE COUNT      131 a    113 c    109 g    142 t      9 others
ORIGIN
```

```
  Query Match             16.0%;  Score 149.8;  DB 10;  Length 504;
  Best Local Similarity   64.6%;  Pred. No. 5.3e-19;
  Matches  239;  Conservative    0;  Mismatches  128;  Indels    3;  Gaps    1;

Qy      25 ttttctcaaccttgattcttagccatggagttctggggaattgaagttaaatcaggaaag 84
           || ||||  ||| | |    | | ||||||||||||||||| |||  || | | ||  |||
Db      36 TTCTCTCTTCCTCGTTCAACAACAATGGAGTTCTGGGGAGTTGCGGTGACACCAAAAAAC 95

Qy      85 ccagttacagtgactcctgaagaaggcattcttatccacgtttctcaggcatcgcttgga 144
             |   ||  ||||||||||||||||  ||  ||| ||||| |||||||||| || |||
Db      96 GCTACTAAGGTGACTCCTGAAGAAGACAGCCTTGTCCACATTTCTCAGGCTTCACTT--- 152

Qy     145 gaatgtaaaaacaagaagggagagtttgtgcctttacatgtaaaggttgggaaccagaac 204
           || || | |   ||   |||| | |||| |||   ||| |  |||||        ||
Db     153 GACTGCACAGTGAAATCTGGAGAATCTGTGGTTTTGAGTGTGACTGTTGGTGGGGCTAAA 212

Qy     205 ttggttctgggaactctatcgactgagaacatccctcagcttttctgtgatttggtattc 264
           | ||| | ||||| || ||    || ||  |||||||| ||  || |||||||| ||
Db     213 CTTGTTATTGGAACACTTTCACAAGACAAGTTCCCTCAGATTAGCTTTGATTTGGTTTTT 272

Qy     265 gacaaggagtttgagctttctcacacttggggaaaaggaagtgtttactttgttggatac 324
           || || ||||||||||||||| ||||   |    |||| || | || | ||  |||| |||
Db     273 GATAAAGAGTTTGAGCTTTCACACAGCGGTACCAAAGCAAATNTTCATTTCATTGGCTAC 332

Qy     325 aaaactcccaacattgagccacaaggctattctgaggaagaagaggaagaagaggaagaa 384
           ||| | |||||||| ||||   | | ||   ||      || || || || |   |||
Db     333 AAATCCCCCAACATCGAGCAGGATGACTTCACTAGTTCGGATGATGAGGATGTTCCTGAA 392

Qy     385 gttcctgctg 394
           | |  | |||
Db     393 GCTGTTCCTG 402


RESULT   7
BF279194
LOCUS       BF279194                 543 bp    mRNA    linear   EST 07-MAR-2001
DEFINITION  GA__Eb0037K03f Gossypium arboreum 7-10 dpa fiber library Gossypium
            arboreum cDNA clone GA__Eb0037K03f, mRNA sequence.
ACCESSION   BF279194
VERSION     BF279194.1  GI:11210264
KEYWORDS    EST.
SOURCE      Gossypium arboreum.
  ORGANISM  Gossypium arboreum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
REFERENCE   1  (bases 1 to 543)
  AUTHORS   Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
            ,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
  TITLE     An integrated analysis of the genetics, development, and evolution
            of the cotton fiber
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: TAATACGACTCACTATAGGG
            High quality sequence stop: 540.
FEATURES             Location/Qualifiers
     source          1. .543
                     /organism="Gossypium arboreum"
                     /strain="AKA"
                     /cultivar="8400"
                     /db_xref="taxon:29729"
                     /clone="GA__Eb0037K03f"
                     /clone_lib="Gossypium arboreum 7-10 dpa fiber library"
                     /tissue_type="Fibers isolated from bolls harvested 7-10
                     dpa"
                     /lab_host="E. coli"
                     /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT      149 a    121 c    121 g    152 t
ORIGIN


  Query Match             15.6%;  Score 146.8;  DB 10;  Length 543;
  Best Local Similarity   60.3%;  Pred. No. 2e-18;
  Matches  288;  Conservative    0;  Mismatches  172;  Indels   18;  Gaps    2;

Qy      18 cctctcttttttctcaaccttgattcttagccatggagttctggggaattgaagttaaatc 77
           ||  || |||| ||   | |  |  | | ||||||||||||||||| ||||||||||  ||
Db      58 CCCTTCCTTTTTTCTTTCGTAGTAATCATCCATGGAGTTCTGGGGTATTGAAGTTAAGTC 117

Qy      78 aggaaagccagttacagtgactcctgaagaaggcattcttatccacgtttctcaggcatc 137
           |||| ||||  ||| ||    |||||  |    |    |||||||  |||| ||||| ||
Db     118 TGGACAGCCTATTAAAGCAGATCCTGGTGCTAACTACGTTATCCATCTTTCCCAGGCTTC 177

Qy     138 gcttggagaatgtaaaaacaagaagggagagtttgtgcctttacatgtaaaggttgggaa 197
           ||||||||  |  || |||   || | ||| |  || ||  |  |||| || ||||
Db     178 ACTTGGAGAGTCCAAGAAC---AAAGCAGAATCCGTTCCACTCTATGTGAATGTTGATGG 234

Qy     198 ccagaacttggttctgggaactctatcgactgagaacatccctcagcttttctgtgattt 257
           ||||  |  |||||||||| || ||      |||    ||||||  | |  | ||||||
Db     235 GAAGAAACTTATTCTGGGAACCCTTTCCCACCAGAGTTGCCCTCAATTATCTTTTGATTT 294

Qy     258 ggtattcgacaaggagtttgagctttctcacacttggggaaaaggaagtgtttacttttgt 317
           |||||  ||| | |||||||||||| || ||||  |||  ||| || || || || || ||  |
Db     295 AGTATTTGACGAAGAGTTTGAGCTCTCCCACAACTGGAAAAACGGGAGCGTATATTTCCT 354

Qy     318 tggatacaaaactcccaacattgagccacaaggctattct--------------gagga 362
           || |||||||||  ||      |||  |   |  ||| |                || ||
Db     355 CGGTTACAAAACTTTCATACCAGAGGAAGGTGATGATTTTGGTATGTCTAGCGAAGAAGA 414

Qy     363 agaagaggaagaagaggaagaagttcctgctgggaatgctgccaaggctgtagctaaacc 422
           |   ||||| || || || ||  | | |||    ||||||    |||| ||  ||
Db     415 AAGCGAGGAGGAGGAAGAGGAGGAGCTTCCTGTTGCTGCTGCTGCTGCTGCAGTAAATGG 474

Qy     423 aaaggctaagcctgcagaagtgaagccagctgttgatgatgaagaggatgagtctgat 480
           ||| ||||||||  |  | |  |  |  | | ||| ||  |  |||  ||| |
Db     475 AAAAGCTAAGCCAGATGGAAAAACAGCCAAGGCTAATGCTGGGAAACCTGATGCTGTT 532


RESULT   8
AV551556
LOCUS       AV551556                 378 bp    mRNA    linear   EST 06-SEP-2000
DEFINITION  AV551556 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
            cDNA clone RZ128c12R 5', mRNA sequence.
ACCESSION   AV551556
VERSION     AV551556.1  GI:8722969
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 378)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     A large scale analysis of cDNA in Arabidopsis thaliana: Generation
            of 12,028 non-redundant expressed sequence tags from normalized and
            size-selected cDNA libraries
  JOURNAL   DNA Res. 7, 175-180 (2000)
  MEDLINE   20363093
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .378
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /clone="RZ128c12R"
                     /clone_lib="Arabidopsis thaliana roots Columbia"
```

```
                     /tissue_type="roots"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT       105 a      84 c      80 g     109 t
ORIGIN


  Query Match             15.5%;  Score 145.8;  DB 9;  Length 378;
  Best Local Similarity   65.6%;  Pred. No. 3.4e-18;
  Matches  229;  Conservative    0;  Mismatches  117;  Indels    3;  Gaps    1;

Qy      25 ttttctcaaccttgattcttagccatggagttctggggaattgaagttaaatcaggaaag 84
           || ||||  ||| | |    | | ||||||||||||||| |||  || | | ||  |||
Db      33 TTCTCTCTTCCTCGTTCAACAACAATGGAGTTCTGGGGAGTTGCGGTGACACCAAAAAAC 92

Qy      85 ccagttacagtgactcctgaagaaggcattcttatccacgtttctcaggcatcgcttgga 144
             |   ||  ||||||||||||||||| ||  ||| ||||| ||||||||| || |||
Db      93 GCTACTAAGGTGACTCCTGAAGAAGACAGCCTTGTCCACATTTCTCAGGCTTCACTT--- 149

Qy     145 gaatgtaaaaacaagaagggagagtttgtgcctttacatgtaaaggttgggaaccagaac 204
           || || | |   ||    ||||| | ||||  |||   ||| |  |||||        ||
Db     150 GACTGCACAGTGAAATCTGGAGAATCTGTGGTTTTGAGTGTGACTGTTGGTGGGGCTAAA 209

Qy     205 ttggttctgggaactctatcgactgagaacatccctcagcttttctgtgatttggtattc 264
             | ||| | ||||| || ||     || ||  ||||||| ||  || |||||||| ||
Db     210 CTTGTTATTGGAACACTTTCACAAGACAAGTTCCCTCAGATTAGCTTTGATTTGGTTTTT 269

Qy     265 gacaaggagtttgagctttctcacacttggggaaaaggaagtgtttactttgttggatac 324
           || || ||||||||||||||| ||||   |   |||| || |||| | ||  |||| |||
Db     270 GATAAAGAGTTTGAGCTTTCACACAGCGGTACCAAAGCAAATGTTCATTTCATTGGCTAC 329

Qy     325 aaaactcccaacattgagccacaaggctattctgaggaagaagaggaag 373
           ||| | ||||||||| ||||   | | ||   ||      || || || |
Db     330 AAATCCCCCAACATCGAGCAGGATGACTTCACTAGTCCGGATGATGAGG 378


RESULT   9
BG131117
LOCUS       BG131117                 759 bp    mRNA    linear   EST 31-JAN-2001
DEFINITION  EST464009 tomato crown gall Lycopersicon esculentum cDNA clone
            cTOE2J5 5' sequence, mRNA sequence.
ACCESSION   BG131117
VERSION     BG131117.1  GI:12631305
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 759)
  AUTHORS   van der Hoeven,R., Sun,H., Cho,J., Utterback,T., Hansen,C., Ronning
            ,C. and Tanksley,S.
  TITLE     Generation of ESTs from tomato crown gall tissue
  JOURNAL   Unpublished (2001)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html.
FEATURES             Location/Qualifiers
     source          1. .759
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cTOE2J5"
                     /clone_lib="tomato crown gall"
                     /tissue_type="crown gall"
                     /dev_stage="crown galls from full-grown plants (8 wks old
                     )"
                     /lab_host="SOLR"
                     /note="Vector: pBluescript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; Four wk old greenhouse plants were stab inoculated
                     on stem with Agrobacterium tumefaciens C58 (Dr. T.J. Burr,
                     Cornell U.). Galls were allowed to develop for another 4
                     wks, when gall tissue was frozen in  liquid nitrogen."
BASE COUNT       246 a     139 c     193 g     181 t
ORIGIN


  Query Match             15.1%;  Score 141.6;  DB 10;  Length 759;
  Best Local Similarity   53.3%;  Pred. No. 1.8e-17;
  Matches  386;  Conservative    0;  Mismatches  284;  Indels   54;  Gaps    2;

Qy      11 taaaaatcctctctttttctcaaccttgattcttagccatggagttctggggaattgaag 70
           |||||  |   | | | ||| || |   |     ||| |||||||| |||||   ||| |
Db      12 TAAAACCCTAATTTCTCTCTAAAACCCTAAAAGCAGCTATGGACTTTTGGGGTGCTGAGG 71

Qy      71 ttaaatcaggaaagccagttacagtgactcctgaagaaggcattcttatccacgtttctc 130
           | |||   ||| ||||| | || ||    |||| ||| || ||  || | ||  |||| |
Db      72 TGAAAAGTGGAGAGCCATTAACTGTACAGCCTGGAGATGGGATGGTTTTGCATCTTTCAC 131

Qy     131 aggcatcgcttggagaatgtaaaaacaagaagggagagtttgtgcctttacatgtaaagg 190
           ||||||| || || || |  || ||  | ||   |||   ||      ||    || ||
Db     132 AGGCATCTCTAGGTGAGTTGAAGAAGGATAAATCAGAAAGTGCCTGCTTGTCAGTGAATA 191

Qy     191 ttgggaaccagaacttggttctgggaactctatcgactgagaacatccctcagcttttct 250
           |||     | ||||  | ||||| || || ||      | |||||  | |||||  |     |
Db     192 TTGATGGCAAGAAACTTGTTCTTGGGACACTCAACTCAGAGAAGGTGCCTCAACAGCAAT 251

Qy     251 gtgatttggtattcgacaaggagtttgagctttctcacacttggggaaaaggaagtgttt 310
           |||  |||| || || || || |||||  | |||||||  | |  ||  || ||||| |
Db     252 TTGACCTGGTTTTTGATAAAGACTTTGAATTATCTCACAACTTGAAAAGTGGCAGTGTCT 311

Qy     311 actttgttggatacaaaaactcccaacattgagccacaaggctattctgaggaagaagagg 370
           ||||| ||||||||||  |  | |||           | |   |   ||| |||||  ||
Db     312 ACTTTTTTGGATACAAGGCAACTAACCCGTTTGATGAGGAAGAGGATGACGAAGATGATT 371

Qy     371 aagaaga----------------------------------------------------gg 379
           | || ||
Db     372 ATGATGATTCTGATGAGGATATCCCCCTCACTCTTGCCAACAGTGGTAAACCTGACCCTA 431

Qy     380 aagaagttcctgctgggaatgctgccaaggctgtagctaaaccaaaggctaagcctgcag 439
           |||| | |          || ||||  |||| |   |||  |  ||||   |||
Db     432 AAGAGGCTGGGAAGTCTAACGCTGGAAAGGATTCTGCTTCAGGAAAGCAGAAGGTGAGGA 491

Qy     440 aagtgaagccagctgttgatgatgaagaggatgagtctgattctgacggaatggatgaag 499
             ||| ||||  ||   |||||||||||| ||    ||||||  ||| ||    ||    |
Db     492 TTGTGGAGCCTACTAAAGATGATGAAGATGAGAGCTCTGATGATGATGGTTCTGACTTGG 551

Qy     500 atgattctgatggtgaggattctgaggaagaagagcctacacctaagaagcctgcatcaa 559
             ||  ||| |  ||   |  ||||   || |||   ||||| ||||| |||    |
Db     552 GAGAGGATGAAGACGACTCTGATGAGAGTGAGGAGGAGACACCAAAGAAGGCTGAGCCTG 611

Qy     560 gcaagaagagagctaatgaaactacccctaaagcacctgtgtcagcaaagaaggcgaaag 619
             || |||||       ||  |  |  | ||| |||||||  | |  || ||||| |||
Db     612 CAAAAAAGAGGAAGGCAGATTCAGCTACAAAAACACCTGTCACTGATAAAAAGGCAAAAT 671

Qy     620 tagcagttactcctcagaaaacagatgagaagaagaaaggggggaaaggctgcaaaccaga 679
           |  ||   || ||||| ||||| |||| || ||    || |   | |  ||||  | |
Db     672 TGACA---ACACCTCAAAAAACTGATGGCAAAAAAGGTGGTGTCCATGTAGCAACACCGC 728

Qy     680 gccc 683
             |||
Db     729 ACCC 732


RESULT  10
BF275467
LOCUS       BF275467                 875 bp    mRNA    linear   EST 07-MAR-2001
DEFINITION  GA__Eb0024C01f Gossypium arboreum 7-10 dpa fiber library Gossypium
            arboreum cDNA clone GA__Eb0024C01f, mRNA sequence.
ACCESSION   BF275467
VERSION     BF275467.1  GI:11206537
KEYWORDS    EST.
SOURCE      Gossypium arboreum.
```

```
  ORGANISM  Gossypium arboreum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
REFERENCE   1  (bases 1 to 875)
  AUTHORS   Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
            ,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
  TITLE     An integrated analysis of the genetics, development, and evolution
            of the cotton fiber
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: TAATACGACTCACTATAGGG
            High quality sequence stop: 630.
FEATURES             Location/Qualifiers
     source          1. .875
                     /organism="Gossypium arboreum"
                     /strain="AKA"
                     /cultivar="8400"
                     /db_xref="taxon:29729"
                     /clone="GA__Eb0024C01f"
                     /clone_lib="Gossypium arboreum 7-10 dpa fiber library"
                     /tissue_type="Fibers isolated from bolls harvested 7-10
                     dpa"
                     /lab_host="E. coli"
                     /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT       265 a    191 c    203 g    216 t
ORIGIN


  Query Match             15.1%;  Score 141.6;  DB 10;  Length 875;
  Best Local Similarity   58.1%;  Pred. No. 1.7e-17;
  Matches  293;  Conservative    0;  Mismatches  199;  Indels   12;  Gaps    2;

Qy      30 tcaaccttgattcttagccatggagttctggggaattgaagttaaatcaggaaagccagt 89
           ||   | | || || |   |||||||||||||| |||||||||||| | ||| ||||  |
Db      79 TCTTTCATCATCCTCATAAATGGAGTTCTGGGGTATTGAAGTTAAAGCTGGACAGCCTAT 138

Qy      90 tacagtgactcctgaagaaggcattcttatccacgtttctcaggcatcgcttggagaatg 149
           || |  |   |||| |   |   | | ||||| |||| |||||| | ||||| || |
Db     139 TAAAACGGACCCTGGTGCCAGTCATGTCATCCATCTTTCCCAGGCAACACTTGGTGAGT- 197

Qy     150 taaaaacaagaagggagagtttgtgcctttacatgtaaaggttgggaaccagaacttggt 209
           ||| || || | ||||| || ||| |  |||| || |||       ||||  | ||
Db     198 --CAAAGAACAAAGCAGAGTCGGTTCCTCTCTATGTGAATGTTAATGGGAAGAAACTCGT 255

Qy     210 tctgggaactctatcgactgagaacatccctcagcttttctgtgatttggtattcgacaa 269
           ||| ||||| || ||    ||||   ||| ||  | |  | |||||| ||||| ||  |
Db     256 TCTCGGAACACTTTCCCACCAGAATTGCCCGCAATTGTGTTTTGATTTAGTATTTGAGCA 315

Qy     270 ggagtttgagctttctcacacttggggaaaaggaagtgtttactttgttggatacaaaac 329
           |  |||||||| |||||||  |||  ||| || || || || ||  |||| |||| |||
Db     316 AGGCTTTGAGCTCTCTCACAACTGGAAAAATGGGAGCGTCTATTTCCTTGGTTACAGAAC 375

Qy     330 tcccaacattgagcca---------caaggctattctgaggaagaagaggaagaagagga 380
           |     |  |||  |          | |  | |  |  || |||||       |||||
Db     376 TTTTGTCCCAGAGGAAGGCTCTGATGATGAATTTGATAGTGAGGAAGAAAGTAGCGAGGA 435

Qy     381 agaagttcctgctgggaatgctgccaaggctgtagctaaaccaaaggctaagcctgcaga 440
           || |  | | ||| |  ||||  || |    ||||||  || | || ||  | ||  |
Db     436 TGAGGAACTTCCTGTGGCTGCTGAAAACGGAAAAGCTAAGACAGATGCAAAAACGGCTAA 495

Qy     441 agtgaagccagctgttgatgatgaagaggatgagtctgattctgacggaatggatgaaga 500
           |    |  | ||| || ||   || | | ||         ||||   ||  ||| | |
Db     496 TGCTGGGAAACCTGATGCTGTAAAACAAGCTGTCAAGATAACTGAACCAAGTGATAACAA 555

Qy     501 tgattctgatggtgaggattctga 524
           |   ||| |  || |||  |||
```

```
Db     556 GAAAAATGAGGACGATGATGATGA 579


RESULT  11
AV549937
LOCUS       AV549937                327 bp    mRNA    linear   EST 06-SEP-2000
DEFINITION  AV549937 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
            cDNA clone RZ106a10R 5', mRNA sequence.
ACCESSION   AV549937
VERSION     AV549937.1  GI:8721350
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 327)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     A large scale analysis of cDNA in Arabidopsis thaliana: Generation
            of 12,028 non-redundant expressed sequence tags from normalized and
            size-selected cDNA libraries
  JOURNAL   DNA Res. 7, 175-180 (2000)
  MEDLINE   20363093
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .327
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /clone="RZ106a10R"
                     /clone_lib="Arabidopsis thaliana roots Columbia"
                     /tissue_type="roots"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT        88 a     74 c     68 g     97 t
ORIGIN


  Query Match             15.0%;  Score 141;  DB 9;  Length 327;
  Best Local Similarity   67.5%;  Pred. No. 2.9e-17;
  Matches  214;  Conservative    0;  Mismatches  100;  Indels    3;  Gaps    1;

Qy      25 ttttctcaaccttgattcttagccatggagttctgggggaattgaagttaaatcaggaaag 84
           || ||||  ||| | |    | | ||||||||||||||| |||  || | | ||  |||
Db      14 TTCTCTCTTCCTCGTTCAACAACAATGGAGTTCTGGGGAGTTGCGGTGACACCAAAAAAC 73

Qy      85 ccagttacagtgactcctgaagaaggcattcttatccacgtttctcaggcatcgcttgga 144
           |   ||  |||||||||||||||| ||  ||| ||||| |||||||||| || |||
Db      74 GCTACTAAGGTGACTCCTGAAGAAGACAGCCTTGTCCACATTTCTCAGGCTTCACTT--- 130

Qy     145 gaatgtaaaaacaagaagggagagtttgtgcctttacatgtaaaggttgggaaccagaac 204
           || || | |   ||    |||||  | ||||  |||   ||| |  |||||       ||
Db     131 GACTGCACAGTGAAATCTGGAGAATCTGTGGTTTTGAGTGTGACTGTTGGTGGGGCTAAA 190

Qy     205 ttggttctgggaactctatcgactgagaacatccctcagcttttctgtgatttggtattc 264
           | ||| | ||||| || ||    || ||  |||||||| ||  || |||||||| ||
Db     191 CTTGTTATTGGAACACTTTCACAAGACAAGTTCCCTCAGATTAGCTTTGATTTGGTTTTT 250

Qy     265 gacaaggagtttgagctttctcacacttggggaaaaggaagtgtttactttgttggatac 324
           || || |||||||||||||| ||||   |    |||| || |||| | ||  |||| |||
Db     251 GATAAAGAGTTTGAGCTTTCACACAGCGGTACCAAAGCAAATGTTCATTTCATTGGCTAC 310

Qy     325 aaaactcccaacattga 341
           ||| | |||||||| ||
Db     311 AAATCCCCCAACATCGA 327


RESULT  12
BI130765
```

```
LOCUS       BI130765                511 bp    mRNA    linear   EST 31-DEC-2001
DEFINITION  G110P49Y Populus cambium cDNA library Populus tremula x Populus
            tremuloides cDNA, mRNA sequence.
ACCESSION   BI130765
VERSION     BI130765.1  GI:18014736
KEYWORDS    EST.
SOURCE      Populus tremula x Populus tremuloides.
  ORGANISM  Populus tremula x Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE   1  (bases 1 to 511)
  AUTHORS   Hertzberg,M., Aspeborg,H., Erlandsson,R., Bjorkbacka,H., Hiltonen
            ,T., Karlsson,J., Teeri,T., Gustafsson,P., Bahlerao,R., Jansson,S.,
            Nilsson,O., Sundberg,B., Nilsson,P., Uhlen,M., Sandberg,G. and
            Lundeberg,J.
  TITLE     Gene expression in Populus
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Erlandsson R
            Department of Biotechnology
            Royal Institute of Technology
            Teknikringen 30, Stockholm S-10044, Sweden
            Tel: 46 8 790 8287
            Fax: 46 8 245452
            Email: rikerl@biochem.kth.se.
FEATURES             Location/Qualifiers
     source          1. .511
                     /organism="Populus tremula x Populus tremuloides"
                     /db_xrei="taxon:47664"
                     /clone_lib="Populus cambium cDNA library"
                     /note="Organ: cambium"
BASE COUNT      137 a    123 c    100 g    151 t
ORIGIN
```

```
  Query Match           14.9%;  Score 140.2;  DB 10;  Length 511;
  Best Local Similarity   64.1%;  Pred. No. 3.7e-17;
  Matches  211;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;

Qy      47 ccatggagttctggggaattgaagttaaatcaggaaagccagttacagtgactcctgaag 106
           ||||||||||||||||  |||||||  ||  | ||  ||||  ||| ||||   ||| |||
Db     155 CCATGGAGTTCTGGGGTGTTGAAGTCAAGGCTGGTGAGCCTCTTAAAGTGGAACCTAAAG 214

Qy     107 aaggcattcttatccacgtttctcaggcatcgcttggagaatgtaaaaacaagaagggag 166
           |   | | | || ||  |||||||||||| | |||||  | | |  ||| |||      |
Db     215 ATTTCTATATGATACATCTTTCTCAGGCAGCCCTTGGTCAGTCTTCAAAGAAGGGAAACG 274

Qy     167 agtttgtgcctttacatgtaaaggttgggaaccagaacttggttctgggaactctatcga 226
           | |  || ||  |     ||||| | |   |  ||||  | ||  ||||||| |||||
Db     275 AATCAGTTCCCCTTTTCCTAAAGTTGGATGAAAAGAAGCTTGTGTTGGGAACCCTATCTC 334

Qy     227 ctgagaacatccctcagcttttctgtgatttggtattcgacaaggagtttgagctttctc 286
           ||||  || |||||||||  | | || ||||||  || || ||  |  ||||| |||||  ||||
Db     335 CTGAAAAGATCCCTCAGTTGTCCTTTGATTTAGTTTTTGAGAGAGAGTTCGAGCTCTCTC 394

Qy     287 acacttggggaaaaggaagtgtttactttgttggatacaaaactcccaacattgagccac 346
           |||  |||  ||| || |||||||  |||   ||| ||| || || | |    |||
Db     395 ACAACTGGAAAAACGGGAGTGTTTTCTTCTGTGGCTACCAAGCTGCTATTCCTGAAAATG 454

Qy     347 aaggctattctgaggaagaagaggaagaa 375
           |     ||| |   || || |  || |||
Db     455 ATTCTGATTTTTCTGATGATGGAGATGAA 483
```

```
RESULT  13
BG445639
LOCUS       BG445639                865 bp    mRNA    linear   EST 15-MAR-2001
DEFINITION  GA__Ea0029A01f Gossypium arboreum 7-10 dpa fiber library Gossypium
            arboreum cDNA clone GA__Ea0029A01f, mRNA sequence.
ACCESSION   BG445639
VERSION     BG445639.1  GI:13355291
KEYWORDS    EST.
SOURCE      Gossypium arboreum.
  ORGANISM  Gossypium arboreum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
REFERENCE   1  (bases 1 to 865)
  AUTHORS   Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
            ,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
  TITLE     An integrated analysis of the genetics, development, and evolution
            of the cotton fiber
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: TAATACGACTCACTATAGGG
            High quality sequence start: 5
            High quality sequence stop: 751.
FEATURES             Location/Qualifiers
     source          1. .865
                     /organism="Gossypium arboreum"
                     /strain="AKA"
                     /cultivar="8400"
                     /db_xref="taxon:29729"
                     /clone="GA__Ea0029A01f"
                     /clone_lib="Gossypium arboreum 7-10 dpa fiber library"
                     /tissue_type="Fibers isolated from bolls harvested 7-10
                     dpa"
                     /lab_host="E. coli"
                     /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT      251 a    170 c    254 g    189 t      1 others
ORIGIN
```

```
  Query Match           14.7%;  Score 137.8;  DB 10;  Length 865;
  Best Local Similarity   53.1%;  Pred. No. 9.1e-17;
  Matches  339;  Conservative    0;  Mismatches  293;  Indels    6;  Gaps    2;

Qy      30 tcaaccttgattcttagccatggagttctggggaattgaagttaaatcaggaaagccagt 89
           ||   | | || || |   |||||||||||||| ||||||||||| | ||| ||||  |
Db      79 TCTTTCATCATCCTCATAAATGGAGTTCTGGGGTATTGAAGTTAAAGCTGGACAGCCTAT 138

Qy      90 tacagtgactcctgaagaaggcattcttatccacgtttctcaggcatcgcttggagaatg 149
           || | |   |||| |   |  | | |||||  |||| |||||| | ||||| || |
Db     139 TAAAACGGACCCTGGTGCCAGTCATGTCATCCATCTTTCCCAGGCAACACTTGGTGAGT- 197

Qy     150 taaaaacaagaagggagagtttgtgcctttacatgtaaaggttgggaaccagaacttggt 209
           |||  || || | |||||  || ||| |  |||| || |||        ||||  | ||
Db     198 --CAAAGAACAAAGCAGAGTCGGTTCCTCTCTATGTGAATGTTAATGGGAAGAAACTCGT 255

Qy     210 tctgggaactctatcgactgagaacatccctcagctttctgtgatttggtattcgacaa 269
           ||| |||||  || ||     ||||   ||| ||  | |  | ||||||  ||||| ||  |
Db     256 TCTCGGAACACTTTCCCACCAGAATTGCCCGCAATTGTGTTTTGATTTAGTATTTGAGCA 315

Qy     270 ggagtttgagctttctcacacttgggaaaaggaagtgtttactttgttggatacaaaac 329
           |  |||||||| |||||||  |||  ||| || || || || ||  |||| |||| |||
Db     316 AGGCTTTGAGCTCTCTCACAACTGGAAAAATGGGAGCGTCTATTTCCTTGGTTACAGAAC 375

Qy     330 tcccaacattgagccacaaggctattctgaggaagaagaggaagaagaggaagaagttcc 389
           |     |  |||    |||||| |  ||| |||   ||    || || |||
Db     376 TTTGTCCCAGAG---GAAGGCTCTGATGATGAATTTGATAGTGAGGAAGAAAGTAGCGA 432

Qy     390 tgctgggaatgctgccaaggctgtagctaaaccaaaggctaagcctgcagaaagtgaagcc 449
           | || | |   | |  |||||  |  ||   ||  |||||| | |  | |   | |
Db     433 GGATGAGGAACTTCCTGTGGCTGCTGAAAACGGAANAGCTAAGACAGATGCAAAAACGGG 492

Qy     450 agctgttgatgatgaagaggatgagtctgattctgacggaatggatgaagatgattctga 509
           || ||   |  || | ||     | ||| |   ||   ||||     |   |
Db     493 TAATGCTGGGAAACCTGATGCTGGAAAACAAGCTGTCAAGATAACTGAACCAAGTGAGTA 552

Qy     510 tggtgaggattctgaggaagaagagcctacacctaagaagcctgcatcaagcaagaagag 569
```

```
Db    553 CAGGAAAAATGAGGACGATGATGATGATAGGGATTCCGAGGATGGAAGTGGGTCTGATGA 612

Qy    570 agctaatgaaactacccctaaagcacctgtgtcagcaaagaaggcgaaagtagcagttac 629

Db    613 TGAAGATGATGCCGAGGATGAAGATGCAGAGATGGCTGGGGAGTGAAGTTCAGATGAGGA 672

Qy    630 tcctcagaaaacagatgagaagaagaaaaggggaaagg 667

Db    673 GGATGAGGAGACACCTAAGAAGGTGGGGCGAGCAAAAG 710


RESULT 14
BF278973
LOCUS       BF278973                517 bp    mRNA    linear   EST 07-MAR-2001
DEFINITION  GA__Eb0036G18f Gossypium arboreum 7-10 dpa fiber library Gossypium
            arboreum cDNA clone GA__Eb0036G18f, mRNA sequence.
ACCESSION   BF278973
VERSION     BF278973.1  GI:11210043
KEYWORDS    EST.
SOURCE      Gossypium arboreum.
  ORGANISM  Gossypium arboreum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
REFERENCE   1  (bases 1 to 517)
  AUTHORS   Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
            ,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
  TITLE     An integrated analysis of the genetics, development, and evolution
            of the cotton fiber
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: TAATACGACTCACTATAGGG
            High quality sequence start: 4
            High quality sequence stop: 514.
FEATURES             Location/Qualifiers
     source          1. .517
                     /organism="Gossypium arboreum"
                     /strain="AKA"
                     /cultivar="8400"
                     /db_xref="taxon:29729"
                     /clone="GA__Eb0036G18f"
                     /clone_lib="Gossypium arboreum 7-10 dpa fiber library"
                     /tissue_type="Fibers isolated from bolls harvested 7-10
                     dpa"
                     /lab_host="E. coli"
                     /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT      158 a    105 c    121 g    133 t
ORIGIN


  Query Match           14.5%;  Score 136.2;  DB 10;  Length 517;
  Best Local Similarity   58.6%;  Pred. No. 2.1e-16;
  Matches  275;  Conservative    0;  Mismatches  188;  Indels    6;  Gaps    2;

Qy      30 tcaaccttgattcttagccatggagttctggggaattgaagttaaatcaggaaagccagt 89

Db      30 TCTTTCATCATCCTCATAAATGGAGTTCTGGGGTATTGAAGTTAAAGCTGGACAGCCTAT 89

Qy      90 tacagtgactcctgaagaaggcattcttatccacgtttctcaggcatcgcttggagaatg 149

Db      90 TAAAACGGACCCTGGTGCCAGTCATGTCATCCATCTTTCCCAGGCAACACTTGGTGAGT- 148

Qy     150 taaaaacaagaagggagagtttgtgcctttacatgtaaaggttgggaaccagaacttggt 209

Db     149 --CAAAGAACAAAGCAGAGTCGGTTCCTCTCTATGTGAATGTTAATGGGAAGAAACTCGT 206

Qy     210 tctgggaactctatcgactgagaacatccctcagcttttctgtgatttggtattcgacaa 269

Db     207 TCTCGGAACACTTTCCCACCAGAATTGCCCGCAATTGTGTTTTGATTTAGTATTTGAGCA 266

Qy     270 ggagtttgagctttctcacacttggggaaaaggaagtgtttactttgttggatacaaaac 329

Db     267 AGGCTTTGAGCTCTCTCACAACTGGAAAAATGGGAGCGTCTATTTCCTTGGTTACAGAAC 326

Qy     330 tcccaacattgagccacaaggctattctgaggaagaagaggaagaagaggaagaagttcc 389

Db     327 TTTTGTCCCAGAG---GAAGGCTCTGATGATGAATTTGATAGTGAGGAAGAAAGTAGCGA 383

Qy     390 tgctgggaatgctgccaaggctgtagctaaaccaaaggctaagcctgcagaagtgaagcc 449

Db     384 GGATGAGGAACTTCCTGTGGCTGCTGAAAACGGAAAAGCTAAGACAGATGCAAAAACGGC 443

Qy     450 agctgttgatgatgaagaggatgagtctgattctgacggaatggatgaa 498

Db     444 TAATGCTGGGAAACCTGATGCTGTAAAACAAGCTGTCAAGATAACTGAA 492


RESULT 15
BF270691
LOCUS       BF270691                715 bp    mRNA    linear   EST 07-MAR-2001
DEFINITION  GA__Eb0009C15f Gossypium arboreum 7-10 dpa fiber library Gossypium
            arboreum cDNA clone GA__Eb0009C15f, mRNA sequence.
ACCESSION   BF270691
VERSION     BF270691.1  GI:11201686
KEYWORDS    EST.
SOURCE      Gossypium arboreum.
  ORGANISM  Gossypium arboreum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
REFERENCE   1  (bases 1 to 715)
  AUTHORS   Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
            ,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
  TITLE     An integrated analysis of the genetics, development, and evolution
            of the cotton fiber
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: TAATACGACTCACTATAGGG
            High quality sequence stop: 665.
FEATURES             Location/Qualifiers
     source          1. .715
                     /organism="Gossypium arboreum"
                     /strain="AKA"
                     /cultivar="8400"
                     /db_xref="taxon:29729"
                     /clone="GA__Eb0009C15f"
                     /clone_lib="Gossypium arboreum 7-10 dpa fiber library"
                     /tissue_type="Fibers isolated from bolls harvested 7-10
                     dpa"
                     /lab_host="E. coli"
                     /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT      212 a    144 c    169 g    190 t
ORIGIN


  Query Match           14.5%;  Score 136;  DB 10;  Length 715;
  Best Local Similarity   63.1%;  Pred. No. 2.1e-16;
  Matches  227;  Conservative    0;  Mismatches  130;  Indels    3;  Gaps    1;

Qy      18 cctctctttttctcaaccttgattcttagccatggagttctggggaattgaagttaaatc 77

Db      78 CCCTTCCTTTTTTCTTTCGTAGTAATCATCCATGGAGTTCTGGGGTATTGAAGTTAAGTC 137
```

```
Qy      78 aggaaagccagttacagtgactcctgaagaaggcattcttatccacgtttctcaggcatc 137
           || ||||  ||| ||    |||||  |    |    |||||||  |||| |||||  ||
Db     138 TAGACAGCCTATTAAAGCAGATCCTGGTGCTAACTACGTTATCCATCTTTCCCAGGCTTC 197

Qy     138 gcttggagaatgtaaaaacaagaagggagagtttgtgcctttacatgtaaaggttgggaa 197
           ||||||||  |  || |||   || | ||| |  || ||  |  |||| || ||||
Db     198 ACTTGGAGAGTCCAAGAAC---AAAGCAGAATCCGTTCCACTCTATGTGAATGTTGATGG 254

Qy     198 ccagaacttggttctgggaactctatcgactgagaacatccctcagcttttctgtgattt 257
           ||||  |  ||||||||||  || ||     |||    ||||||  | |  | ||||||
Db     255 GAAGAAACTTATTCTGGGAACCCTTTCCCACCAGAGTTGCCCTCAATTATCTTTTGATTT 314

Qy     258 ggtattcgacaaggagtttgagctttctcacacttggggaaaaggaagtgtttacttttgt 317
           |||||  |||  |  ||||||||||| || ||||  |||  ||| || || || || ||  |
Db     315 AGTATTTGACGAAGAGTTTGAGCTCTCCCACAACTGGAAAAACGGGAGCGTATATTTCCT 374

Qy     318 tggatacaaaactcccaacattgagccacaaggctattctgaggaagaagaggaaagaaga 377
           || |||||||||  ||     |||  |   |   ||| ||           ||||||||
Db     375 CGGTTACAAAACTTTCATACCAGAGGAAGGTGATGATTTTGGTATGTCTAGCGAAGAAGA 434
```

Search completed: May 2, 2002, 18:43:44
Job time: 10868 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: May 2, 2002, 20:14:52 ; Search time 144.83 Seconds
(without alignments)
187.897 Million cell updates/sec

Title: US-09-645-337-6
Perfect score: 1279
Sequence: 1 MEFWGIEVKSGKPVTVTPEE.....TFNSGNALESHNKAKHAAAK 245

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1279	100.0	245	21	AAG38221	Arabidopsis thalia
2	1279	100.0	245	22	AAB67813	Amino acid sequenc
3	1274	99.6	257	21	AAG31484	Arabidopsis thalia
4	1261.5	98.6	248	21	AAG10894	Arabidopsis thalia
5	606	47.4	305	21	AAG26185	Arabidopsis thalia
6	594	46.4	305	22	AAB67814	Amino acid sequenc
7	513.5	40.1	311	20	AAY28804	Maize histone deac
8	504	39.4	98	21	AAG15354	Arabidopsis thalia
9	499	39.0	302	20	AAY28803	Maize histone deac
10	487	38.1	305	20	AAY28802	Maize histone deac
11	446	34.9	285	20	AAY28805	Maize histone deac
12	405	31.7	126	21	AAG15353	Arabidopsis thalia
13	374.5	29.3	208	21	AAG25823	Arabidopsis thalia
14	306	23.9	203	21	AAG09171	Arabidopsis thalia
15	255.5	20.0	155	21	AAG17704	Arabidopsis thalia
16	241	18.8	194	21	AAG54925	Arabidopsis thalia
17	230	18.0	181	21	AAG09172	Arabidopsis thalia
18	224	17.5	412	20	AAW68010	S. frugiperda immu
19	203.5	15.9	103	21	AAG49038	Arabidopsis thalia
20	183.5	14.3	147	21	AAG09173	Arabidopsis thalia
21	175.5	13.7	349	22	ABB62190	Drosophila melanog
22	155	12.1	208	22	ABG14460	Novel human diagno
23	155	12.1	277	22	ABG05501	Novel human diagno
24	155	12.1	317	21	AAB28184	Human Nucleophosmi
25	155	12.1	334	22	ABG12631	Novel human diagno
26	153	12.0	536	22	AAB92836	Human protein sequ
27	153	12.0	707	16	AAR79912	Human nucleolin.
28	153	12.0	707	20	AAW84052	Human V3 loop HIV
29	153	12.0	707	22	AAB48964	Human nucleolin.
30	150	11.7	330	22	ABG07612	Novel human diagno
31	150	11.7	338	22	ABG17097	Novel human diagno
32	147.5	11.5	362	22	ABG19618	Novel human diagno
33	140.5	11.0	268	21	AAG13510	Arabidopsis thalia
34	140.5	11.0	268	21	AAG37328	Arabidopsis thalia
35	140.5	11.0	310	21	AAG13509	Arabidopsis thalia
36	140.5	11.0	310	21	AAG37327	Arabidopsis thalia
37	140.5	11.0	450	21	AAG38748	Arabidopsis thalia
38	140.5	11.0	484	21	AAG38747	Arabidopsis thalia
39	140.5	11.0	590	21	AAG38746	Arabidopsis thalia
40	140.5	11.0	612	21	AAG35761	Arabidopsis thalia
41	140.5	11.0	622	21	AAG35760	Arabidopsis thalia
42	140.5	11.0	649	21	AAG35759	Arabidopsis thalia
43	140.5	11.0	1881	21	AAY44506	Streptococcus pneu
44	140	10.9	244	20	AAY27876	Human secreted pro
45	139.5	10.9	411	20	AAW68011	Yeast immunophilin

ALIGNMENTS

RESULT 1
AAG38221
ID AAG38221 standard; Protein; 245 AA.
XX
AC AAG38221;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 47121.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.

```
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match           100.0%;  Score 1279;  DB 21;  Length 245;
  Best Local Similarity  100.0%;  Pred. No. 2.1e-107;
  Matches  245;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mefwgievksgkpvtvtpeegilihvsqaslgecknkkgefvplhvkvgnqnlvlgtlst 60

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGNAAK 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 enipqlfcdlvfdkefelshtwgkgsvyfvgyktpniepqgyseeeeeeeeevpagnaak 120

Qy     121 AVAKPKAKPAEVKPAVDDEEDESDSDGMDEDDSDGEDSEEEEPTPKKPASSKKRANETTP 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 avakpkakpaevkpavddeedesdsdgmdeddsdgedseeeeptpkkpasskkranettp 180

Qy     181 KAPVSAKKAKVAVTPQKTDEKKKGGKAANQSPKSASQVSCGSCKKTFNSGNALESHNKAK 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 kapvsakkakvavtpqktdekkkggkaanqspksasqvscgsckktfnsgnaleshnkak 240

Qy     241 HAAAK 245
           |||||
Db     241 haaak 245


RESULT   2
AAB67813
ID   AAB67813 standard; Protein; 245 AA.
XX
AC   AAB67813;
XX
DT   29-JUN-2001  (first entry)
XX
DE   Amino acid sequence of a histone deacetylase designated AtHDA2A.
XX
KW   Histone deacetylase; AtHDA2A; RPD3; gene expression; transgenic plant;
KW   HDA1; ethylene-responsive phenotype; hypocotyl elongation.
XX
OS   Arabidopsis thaliana.
XX
PN   CA2316036-A1.
XX
PD   27-FEB-2001.
XX
PF   24-AUG-2000; 2000CA-2316036.
XX
PR   27-AUG-1999;   99US-0383971.
XX
PA   (MIAC ) CANADA MIN AGRIC & AGRI-FOOD CANADA.
XX
PI   Miki B,  Brown D,  Tian L,  Wu K;
XX
DR   WPI; 2001-258457/27.
DR   N-PSDB; AAF80352.
XX
PT   Methods for regulating gene expression in transgenic plants, e.g.
PT   repressing ethylene-responsive phenotypes (e.g. inhibition of hypocotyl
PT   elongation), comprises introducing genes encoding histone deacetylase
PT   -
XX
PS   Claim 10; Fig 2A; 91pp; English.
XX
CC   The present sequence represents Arabidopsis thaliana histone deacetylase
CC   designated AtHDA2A. The protein is homologous to yeast RPD3 and HDA1.
CC   The polynucleotide sequence is used in the method of the invention.
CC   The specification describes a method for regulating gene expression in
CC   transgenic plants. The method comprises modifying histones by introducing
CC   chimeric nucleotide sequences which have regulatory elements in operative
CC   association with a gene of interest or with a nucleotide sequence
CC   encoding histone deacetylase. The method is useful for regulating the
CC   developmental, physiological or biochemical pathway within a plant,
CC   particularly for repressing ethylene-responsive phenotypes
CC   (e.g. inhibition of hypocotyl elongation). The method is also useful
CC   as a functional test for identifying a phenotype associated with
CC   perturbing a gene. The histone deacetylase genes are useful for
CC   altering the development of an organism.
XX
SQ   Sequence   245 AA;


  Query Match           100.0%;  Score 1279;  DB 22;  Length 245;
  Best Local Similarity  100.0%;  Pred. No. 2.1e-107;
  Matches  245;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mefwgievksgkpvtvtpeegilihvsqaslgecknkkgefvplhvkvgnqnlvlgtlst 60

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGNAAK 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 enipqlfcdlvfdkefelshtwgkgsvyfvgyktpniepqgyseeeeeeeeevpagnaak 120

Qy     121 AVAKPKAKPAEVKPAVDDEEDESDSDGMDEDDSDGEDSEEEEPTPKKPASSKKRANETTP 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 avakpkakpaevkpavddeedesdsdgmdeddsdgedseeeeptpkkpasskkranettp 180

Qy     181 KAPVSAKKAKVAVTPQKTDEKKKGGKAANQSPKSASQVSCGSCKKTFNSGNALESHNKAK 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 kapvsakkakvavtpqktdekkkggkaanqspksasqvscgsckktfnsgnaleshnkak 240

Qy     241 HAAAK 245
           |||||
Db     241 haaak 245


RESULT   3
AAG31484
ID   AAG31484 standard; Protein; 257 AA.
XX
AC   AAG31484;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 37817.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
```

PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.

```
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match             99.6%;  Score 1274;  DB 21;  Length 257;
  Best Local Similarity   100.0%;  Pred. No. 6.3e-107;
  Matches  244;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mefwgievksgkpvtvtpeegilihvsqaslgecknkkgefvplhvkvgnqnlvlgtlst 60

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGNAAK 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 enipqlfcdlvfdkefelshtwgkgsvyfvgyktpniepqgyseeeeeeeeevpagnaak 120

Qy     121 AVAKPKAKPAEVKPAVDDEEDESDSDGMDEDDSDGEDSEEEEPTPKKPASSKKRANETTP 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 avakpkakpaevkpavddeedesdsdgmdeddsdgedseeeeptpkkpasskkranettp 180

Qy     181 KAPVSAKKAKVAVTPQKTDEKKKGGKAANQSPKSASQVSCGSCKKTFNSGNALESHNKAK 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 kapvsakkakvavtpqktdekkkggkaanqspksasqvscgsckktfnsgnaleshnkak 240

Qy     241 HAAA 244
           ||||
Db     241 haaa 244


RESULT   4
AAG10894
ID   AAG10894 standard; Protein; 248 AA.
XX
AC   AAG10894;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 9401.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
```

PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 98.6%; Score 1261.5; DB 21; Length 248;
Best Local Similarity 98.4%; Pred. No. 8e-106;
Matches 244; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

```
Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mefwgievksgkpvtvtpeegililhvsqaslgecknkkgefvplhvkvgnqnlvlgtlst 60

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYS---EEEEEEEEEVPAGN 117
           ||||||||||||||||||||||||||||||||||||||||||   ||||||||||||||
Db      61 enipqlfcdlvfdkefelshtwgkgsvyfvgyktpniepqgyseeeeeeeeeeevpagn 120

Qy     118 AAKAVAKPKAKPAEVKPAVDDEEDESDSDGMDEDDSDGEDSEEEEPTPKKPASSKKRANE 177
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 aakavakpkakpaevkpavddeedesdsdgmdeddsdgedseeeeptpkkpasskkrane 180

Qy     178 TTPKAPVSAKKAKVAVTPQKTDEKKKGGKAANQSPKSASQVSCGSCKKTFNSGNALESHN 237
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db     181 ttpkapvsakkakvavtpqktdekkkggkaanxspksasqvscgsckktfnsgnaleshn 240

Qy     238 KAKHAAAK 245
           ||||||||
Db     241 kakhaaak 248
```

```
RESULT   5
AAG26185
ID   AAG26185 standard; Protein; 305 AA.
XX
AC   AAG26185;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 30547.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
PR   02-AUG-1999;   99US-0146386.
PR   02-AUG-1999;   99US-0146388.
PR   02-AUG-1999;   99US-0146389.
PR   03-AUG-1999;   99US-0147038.
PR   04-AUG-1999;   99US-0147204.
PR   04-AUG-1999;   99US-0147302.
PR   05-AUG-1999;   99US-0147192.
PR   05-AUG-1999;   99US-0147260.
PR   06-AUG-1999;   99US-0147303.
PR   06-AUG-1999;   99US-0147416.
PR   09-AUG-1999;   99US-0147493.
PR   09-AUG-1999;   99US-0147935.
PR   10-AUG-1999;   99US-0148171.
PR   11-AUG-1999;   99US-0148319.
PR   12-AUG-1999;   99US-0148341.
PR   13-AUG-1999;   99US-0148565.
PR   13-AUG-1999;   99US-0148684.
PR   16-AUG-1999;   99US-0149368.
PR   17-AUG-1999;   99US-0149175.
PR   18-AUG-1999;   99US-0149426.
PR   20-AUG-1999;   99US-0149722.
PR   20-AUG-1999;   99US-0149723.
```

```
PR   20-AUG-1999;   99US-0149929.
PR   23-AUG-1999;   99US-0149902.
PR   23-AUG-1999;   99US-0149930.
PR   25-AUG-1999;   99US-0150566.
PR   26-AUG-1999;   99US-0150884.
PR   27-AUG-1999;   99US-0151065.
PR   27-AUG-1999;   99US-0151066.
PR   27-AUG-1999;   99US-0151080.
PR   30-AUG-1999;   99US-0151303.
PR   31-AUG-1999;   99US-0151438.
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match             47.4%;  Score 606;  DB 21;  Length 305;
  Best Local Similarity   50.3%;  Pred. No. 1.5e-46;
  Matches  150;  Conservative   24;  Mismatches   66;  Indels   58;  Gaps   13;

Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
           |||||: |       |||||  |:|:||||| :|  | || | | | ||   ||:||||
Db       1 mefwgvavtpknatkvtpeedslvhisqasl-dctvksgesvvlsvtvggaklvigtlsq 59

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGY-SEEEEEEEEEVPA---- 115
             :  ||:  ||||||||||||:  | :|:|:|||:||||   : | ::|:  | |||
Db      60 dkfpqisfdlvfdkefelshsqtkanvhfiqykspnieqddftssddedvpeavpapapt 119

Qy     116 -----GNAAKAV----AKPKAKPAEVKPAVD----DEEDESDSD----------GM--D 149
                 |||  ||    :|||||||||||| :    ||||||| :           ||  |
Db     120 avtangnaqaavvkadskpkakpaevkpaeekpesdeedesddedeseedddsekqmdvd 179

Qy     150 EDDSDG---EDS--EEEEPTPKKPAS-SKKRANETTPKAPVSAKKAKVAV----TPQKTD 199
           |||||    |||  |||| |||||   :||| ||:  | ||| |||| |     |||||:
Db     180 eddsdddeeedsedeeeeetpkkpepinkkrpnesvsktpvsgkkakpaaapastpqkte 239

Qy     200 EKKKGGKA--------------ANQSPKSASQVSCG-SCKKTPNSGNALESHNKAKH 241
           ||||||                |||||||  | | | : || ||||      ||  :
Db     240 ekkkqghtatphpakkggkspvnanqspksqqqssqqnnnkkpfnsqkqfqqsnkqsn 297


RESULT   6
AAB67814
ID   AAB67814 standard; Protein; 305 AA.
XX
AC   AAB67814;
XX
DT   29-JUN-2001  (first entry)
XX
DE   Amino acid sequence of a histone deacetylase designated AtHDA2B.
XX
KW   Histone deacetylase; AtHDA2B; RPD3; gene expression; transgenic plant;
KW   HDA1; ethylene-responsive phenotype; hypocotyl elongation.
XX
OS   Arabidopsis thaliana.
XX
PN   CA2316036-A1.
XX
PD   27-FEB-2001.
XX
PF   24-AUG-2000; 2000CA-2316036.
XX
PR   27-AUG-1999;   99US-0383971.
XX
PA   (MIAC ) CANADA MIN AGRIC & AGRI-FOOD CANADA.
XX
PI   Miki B,  Brown D,  Tian L,  Wu K;
XX
DR   WPI; 2001-258457/27.
DR   N-PSDB; AAF80353.
XX
PT   Methods for regulating gene expression in transgenic plants, e.g.
PT   repressing ethylene-responsive phenotypes (e.g. inhibition of hypocotyl
PT   elongation), comprises introducing genes encoding histone deacetylase
PT   -
XX
PS   Claim 10; Fig 2B; 91pp; English.
XX
CC   The present sequence represents Arabidopsis thaliana histone deacetylase
CC   designated AtHDA2B. The protein is homologous to yeast RPD3 and HDA1.
CC   The polynucleotide sequence is used in the method of the invention.
CC   The specification describes a method for regulating gene expression in
CC   transgenic plants. The method comprises modifying histones by introducing
CC   chimeric nucleotide sequences which have regulatory elements in operative
CC   association with a gene of interest or with a nucleotide sequence
CC   encoding histone deacetylase. The method is useful for regulating the
CC   developmental, physiological or biochemical pathway within a plant,
CC   particularly for repressing ethylene-responsive phenotypes
CC   (e.g. inhibition of hypocotyl elongation). The method is also useful
CC   as a functional test for identifying a phenotype associated with
CC   perturbing a gene. The histone deacetylase genes are useful for
CC   altering the development of an organism.
XX
SQ   Sequence   305 AA;


  Query Match             46.4%;  Score 594;  DB 22;  Length 305;
  Best Local Similarity   50.7%;  Pred. No. 1.8e-45;
  Matches  151;  Conservative   26;  Mismatches   63;  Indels   58;  Gaps   15;

Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
           |||||: |       |||||  |:|:||||| :|  | || | | | ||   ||:||||
Db       1 mefwgvavtpknatkvtpeedslvhisqasl-dctvksgesvvlsvtvggaklvigtlsq 59

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGY-SEEEEEEEEEVPA---- 115
```

```
          :  ||:  ||||||||||||:  | :|:|:|||:||||   : | ::|:  | |||
Db     60 dkfpqisfdlvfdkefelshsgtkanvhfigykspnieqddftssddedvpeavpapapt 119

Qy    116 -----GNAAKAV----AKPKAKPAEVKPAVD----DEEDESDSD----------GM--D 149
              |||  ||     ||||||||||||| :    ||||||| :           ||  |
Db    120 avtangnagaavvkadtkpkakpaevkpaeekpesdeedesddedeseedddsekgmdvd 179

Qy    150 EDDSDG---EDS--EEEEPTPKKPAS-SKKRANETTPKAPVSAKKAKVAV----TPQKTD 199
          |||||    |||  |||| |||||   :||| ||:  | ||| |||| |     |||||:
Db    180 eddsdddeeedsedeeeeetpkkpepinkkrpnesvsktpvsgkkakpaaapastpqkte 239

Qy    200 EK----------KKGGKA---ANQSPKSASQVSCG-SCKKTFNSGNAL-ESHNKAKH 241
          :|          |||||:   |||||||  | | | : || ||||     |:||  :
Db    240 kkkgghtatphpakkggkspvnanqspksggqssggnnnkkpfnsgkqfggsnnkgsn 297
```

RESULT 7
AAY28804
ID AAY28804 standard; protein; 311 AA.
XX
AC AAY28804;
XX
DT 13-JAN-2000 (first entry)
XX
DE Maize histone deacetylase-8.
XX
KW Maize histone deacetylase; HD; HD cDNA; family 2, ZmHD2; gene repression;
KW acetyl modification; promoter; regulatory element; transgenic plant;
KW disease resistance; toxin screening; pathogenicity;
KW disease response promoter.
XX
OS Zea mays.
XX
PN WO9951731-A2.
XX
PD 14-OCT-1999.
XX
PF 02-APR-1999; 99WO-US07370.
XX
PR 03-APR-1998; 98US-0080563.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Baldwin DA, Briggs SP, Crane VC;
XX
DR WPI; 1999-611038/52.
DR N-PSDB; AAX90844.
XX
PT New deacetylase genes, used for producing transgenic plants which have
PT increased disease resistance -
XX
PS Claim 1; Page 79-80; 87pp; English.
XX
CC The present sequence is maize histone deacetylase encoded by HD cDNA
CC belonging to family 2, ZmHD2. This enzyme responsible for removing acetyl
CC modifications, may be localised to promoters targeted for repression by
CC other proteins that associate with HD and specifically bind regulatory
CC elements in promoter DNA. The HD nucleotide sequence can be used for
CC producing transgenic plants with increased disease resistance.
CC Additionally, compositions find use in screening for toxins that affect
CC pathogenicity and in determining which disease response promoters are
CC regulated by histone deacetylase.
XX
SQ Sequence 311 AA;

Query Match 40.1%; Score 513.5; DB 20; Length 311;
Best Local Similarity 42.3%; Pred. No. 3.3e-38;
Matches 127; Conservative 29; Mismatches 81; Indels 63; Gaps 9;

```
Qy      1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
          |||||:||| |  |   |  | ::|||||:|||  :|| :   ::||| :: | :||||
Db      1 mefwglevkpgstvkcepghgfilhvsqaalge--skksdsalmyvkvddkklaigtlsi 58

Qy     61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPN1---EPQGYSEEEEEEEEVPA-- 115
          :  ||:  ||||:|||||||    ||:| |||       :      |:|||| :|
Db     59 dkypqiqfdlvfnkefelshtskttsvffsgykveqpiegdemdldsedeeeelnipvik 118

Qy    116 ----------GNAAKAVAKPKAKPA---EVKPAVD-DEEDESDSDGMDEDDSD------ 154
                     |  ||||   :| :   | |   | |: || :|| ||||||
Db    119 engkadgkeeqknqekavaataskssiglekkskddsddsdedesddsdeddsddsdege 178

Qy    155 -----------------GEDSEEEEPTPKKPASSKKRANETTPKAPVSAKKAKVAVTP-Q 196
                          :| ||| |||||| : |||  |   | |:| ||||||  | |
Db    179 glspdegdddssdeddtsdddeeetptpkkpeagkkrgaenalktplsdkkakvatppaq 238

Qy    197 KTDEKK----------KGGKAAN------QSPKSASQVSCGSCKKTFNSGNALESHNKA 239
          ||  ||           ||   ||      :||||   | | || |||||  ||::|:||
Db    239 ktggkkgathvatphpakgktpanndkstekspksggsvpckscsktfnsemalqahska 298
```

RESULT 8
AAG15354
ID AAG15354 standard; Protein; 98 AA.
XX
AC AAG15354;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 15575.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.

```
PR   29-OCT-1999;   99US-0162142.

  Query Match             39.4%;  Score 504;  DB 21;  Length 98;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-38;
  Matches   98;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    148 MDEDDSDGEDSEEEEPTPKKPASSKKRANETTPKAPVSAKKAKVAVTPQKTDEKKKGGKA 207
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 mdeddsdgedseeeeptpkkpasskkranettpkapvsakkakvavtpqktdekkkggka 60

Qy    208 ANQSPKSASQVSCGSCKKTFNSGNALESHNKAKHAAAK 245
          |||||||||||||||||||||||||||||||||||||
Db     61 anqspksasqvscgsckktfnsgnaleshnkakhaaak 98


RESULT   9
AAY28803
ID   AAY28803 standard; protein; 302 AA.
XX
AC   AAY28803;
XX
DT   13-JAN-2000  (first entry)
XX
DE   Maize histone deacetylase-7.
XX
KW   Maize histone deacetylase; HD; HD cDNA; family 2, ZmHD2; gene repression;
KW   acetyl modification; promoter; regulatory element; transgenic plant;
KW   disease resistance; toxin screening; pathogenicity;
KW   disease response promoter.
XX
OS   Zea mays.
XX
PN   WO9951731-A2.
XX
PD   14-OCT-1999.
XX
PF   02-APR-1999;   99WO-US07370.
XX
PR   03-APR-1998;   98US-0080563.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Baldwin DA,  Briggs SP,  Crane VC;
XX
DR   WPI; 1999-611038/52.
DR   N-PSDB; AAX90843.
XX
PT   New deacetylase genes, used for producing transgenic plants which have
PT   increased disease resistance  -
XX
PS   Claim 1; Page 76-77; 87pp; English.
XX
CC   The present sequence is maize histone deacetylase encoded by HD cDNA
CC   belonging to family 2, ZmHD2. This enzyme responsible for removing acetyl
CC   modifications, may be localised to promoters targeted for repression by
CC   other proteins that associate with HD and specifically bind regulatory
CC   elements in promoter DNA. The HD nucleotide sequence can be used for
CC   producing transgenic plants with increased disease resistance.
CC   Additionally, compositions find use in screening for toxins that affect
CC   pathogenicity and in determining which disease response promoters are
CC   regulated by histone deacetylase.
XX
SQ   Sequence   302 AA;


  Query Match             39.0%;  Score 499;  DB 20;  Length 302;
  Best Local Similarity   40.5%;  Pred. No. 6.5e-37;
  Matches  124;  Conservative   36;  Mismatches   74;  Indels   72;  Gaps   11;

Qy      1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
          |||||:||| |  |   | ::|:|||:|||  :|| :   ::||: :| | :||||
Db      1 mefwglevkpqstvkcepgygfvlhlsqaalge--skksdnalmyvkiddqklaigtlsv 58

Qy     61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSE-------EEEEEEEEV 113
          :  | :  ||:|||||||||||    ||:| |||      | : |       |:|:||  |
Db     59 dknphiqfdlifdkefelshtskttsvfftgykv----eqpfeedemdldsededeelnv 114

Qy    114 PA----GNA--------AKAVAKP-KAKPAEVKPAVDDEEDESDSDGMDEDDSD------ 154
          |     | |          |||| | |: |   |   ||: || ::|  |||::|
Db    115 pvvkengkadgkkqksqekavaapsksspdskkskddddsdedetddsdedetddsdegl 174

Qy    155 --------------GEDSEEEEPTPKKPASSKKRANETTP-KAPVSAKKAKVAVTP--Q 196
                        :| ||: ||||||   |||| |::  | |:| |||||| ||  |
Db    175 sseegdddssdeddtsddeeedtptpkkpevgkkraaessvlktplsdkkakva-tpssq 233

Qy    197 KTDEKK--------------------KGGKAANQSPKSASQVSCGSCKKTFNSGNALES 235
          ||  ||                    |  |:   :|||   | | || |:| |  ||::
Db    234 ktggkkgaavhvatphpakgktivnndksvkspksapksggsvpckscsksfisetalqa 293

Qy    236 HNKAKH 241
          |:||||
Db    294 hskakh 299


RESULT  10
AAY28802
ID   AAY28802 standard; protein; 305 AA.
XX
AC   AAY28802;
XX
DT   13-JAN-2000  (first entry)
XX
DE   Maize histone deacetylase-6.
XX
KW   Maize histone deacetylase; HD; HD cDNA; family 2, ZmHD2; gene repression;
KW   acetyl modification; promoter; regulatory element; transgenic plant;
KW   disease resistance; toxin screening; pathogenicity;
KW   disease response promoter.
XX
OS   Zea mays.
XX
PN   WO9951731-A2.
XX
PD   14-OCT-1999.
XX
PF   02-APR-1999;   99WO-US07370.
XX
PR   03-APR-1998;   98US-0080563.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Baldwin DA,  Briggs SP,  Crane VC;
XX
DR   WPI; 1999-611038/52.
DR   N-PSDB; AAX90842.
XX
PT   New deacetylase genes, used for producing transgenic plants which have
PT   increased disease resistance  -
XX
PS   Claim 1; Page 72-73; 87pp; English.
XX
CC   The present sequence is maize histone deacetylase encoded by HD cDNA
CC   belonging to family 2, ZmHD2. This enzyme responsible for removing acetyl
CC   modifications, may be localised to promoters targeted for repression by
CC   other proteins that associate with HD and specifically bind regulatory
CC   elements in promoter DNA. The HD nucleotide sequence can be used for
CC   producing transgenic plants with increased disease resistance.
CC   Additionally, compositions find use in screening for toxins that affect
CC   pathogenicity and in determining which disease response promoters are
CC   regulated by histone deacetylase.
XX
SQ   Sequence   305 AA;


  Query Match             38.1%;  Score 487;  DB 20;  Length 305;
  Best Local Similarity   41.2%;  Pred. No. 8e-36;
```

```
  Matches  127;  Conservative   37;  Mismatches   74;  Indels   70;  Gaps   13;

Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
           |||||:||| |   |    |  | ::|:|||:|||  :|| :   ::||: :| | :||||
Db       1 mefwglevkpqstvkcepgygfvlhlsqaalqe--skksdnalmyvkiddqklaigtlsv 58

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSE-------EEEEEEEEV 113
           :  | :  ||:|||||||||||    ||:| |||       | : |       |:|:||  |
Db      59 dknphiqfdllfdkefelshtsktlsvfftqykv----eqpfeedemdldsededeelnv 114

Qy     114 PA----GNA--------AKAVAKP-KAKPAEVKPAVDDEEDESDSDGMDEDDSD------ 154
           |     | |         |||| | |: |   |   ||: || ::|  |||::|
Db     115 pvvkengkadekkqksqekavaapsksspdskkskddddsdedetddsdedetddsdegl 174

Qy     155 --------------GEDSEEEEPTPKKPASSKKRANETTP-KAPVSAKKAKVAVTP--Q 196
                         :| ||: ||||||   |||| |::  | |:| ||||||  ||  |
Db     175 speegdddssdeddtsddeeedtptpkkpevgkkraaessvlktplsdkkakva-tpssq 233

Qy     197 KTDEKK------------KGGKAAN-----QSPKSA--SQVSCGSCKKTFNSGNALESHN 237
           ||  ||            ||    |     :|||||  | | | || |:| |  ||::|:
Db     234 ktggkkqaavhvatphpakqktivnndksvkspksapksqvpckscsksfisetalqahs 293

Qy     238 KAKHAAAK 245
           ||   |::
Db     294 kanmgase 301


RESULT  11
AAY28805
ID   AAY28805 standard; protein; 285 AA.
XX
AC   AAY28805;
XX
DT   13-JAN-2000  (first entry)
XX
DE   Maize histone deacetylase-9.
XX
KW   Maize histone deacetylase; HD; HD cDNA; family 2, ZmHD2; gene repression;
KW   acetyl modification; promoter; regulatory element; transgenic plant;
KW   disease resistance; toxin screening; pathogenicity;
KW   disease response promoter.
XX
OS   Zea mays.
XX
PN   WO9951731-A2.
XX
PD   14-OCT-1999.
XX
PF   02-APR-1999;   99WO-US07370.
XX
PR   03-APR-1998;   98US-0080563.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Baldwin DA,  Briggs SP,  Crane VC;
XX
DR   WPI; 1999-611038/52.
DR   N-PSDB; AAX90845.
XX
PT   New deacetylase genes, used for producing transgenic plants which have
PT   increased disease resistance  -
XX
PS   Claim 1; Page 82-84; 87pp; English.
XX
CC   The present sequence is maize histone deacetylase encoded by HD cDNA
CC   belonging to family 2, ZmHD2. This enzyme responsible for removing acetyl
CC   modifications, may be localised to promoters targeted for repression by
CC   other proteins that associate with HD and specifically bind regulatory
CC   elements in promoter DNA. The HD nucleotide sequence can be used for
CC   producing transgenic plants with increased disease resistance.
CC   Additionally, compositions find use in screening for toxins that affect
CC   pathogenicity and in determining which disease response promoters are
CC   regulated by histone deacetylase.
XX
SQ   Sequence   285 AA;


  Query Match             34.9%;  Score 446;  DB 20;  Length 285;
  Best Local Similarity   41.9%;  Pred. No. 3.7e-32;
  Matches  122;  Conservative   31;  Mismatches   84;  Indels   54;  Gaps   12;

Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
           ||||| ||| |  |:    :|::||:|||:||| | |  |   : ||: :: |||||||
Db       1 mefwgeevkpgatvsckvgdglvihlsqaalqepk-kasenaivsvkiddkklvlqtlsv 59

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTP----NIEPQGYSEEEEEEEEVPAG 116
           |  ||: |||||||:|||||     ||:| |||:|      :    | :|| | : :|
Db      60 ekhpqiscdlvfdkdfelshnsktasvffcgykspvplfesdsqedssdeevepdlipmq 119

Qy     117 N------AAKAVAK---PKAKPAEVKPAVDDEEDESDSDGMDEDDSDGED---------- 157
           |        ||  |     |  |     ||:  :|||:  :||:|  ||
Db     120 nneikistakvpvkvgiqnadedetssgddddftdsdsemseedessdedevssdtdtsd 179

Qy     158 ---SEEEEPTPKKP--ASSKKRANET-TPKAPVSAKKAKVAVTPQKTDEK-------KKG 204
               |||: |||||      |||| |  ||    | ||||   : ||| :|       |:
Db     180 dsgseeqtptpkktevvvgkkraieaetp----sgkkakseqsaqktqdkvstshpakqs 235

Qy     205 GK-----------AANQSPKSASQVSCGSCKKTFNSGNALESHNKA-KHAA 243
           |           |  :|||| |  :| || |:| | :|||| || || |
Db     236 sktpadkstktptadkkspksgsh-ackscsksfgsasaleshqkakkhea 285


RESULT  12
AAG15353
ID   AAG15353 standard; Protein; 126 AA.
XX
AC   AAG15353;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 15574.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
```

PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.

PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 31.7%; Score 405; DB 21; Length 126;
Best Local Similarity 71.7%; Pred. No. 6.5e-29;
Matches 81; Conservative 11; Mismatches 19; Indels 2; Gaps 2;

```
Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mefwglevksgkpvtvtpeegilihvsqaslgecknkkgefvplhvkvgnqnlvlgtlst 60

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVY-FVGYKTPNIEPQGYSEEEEEEEEE 112
           ||||||||||||||    | ||:  |:  :  | | :  :    :: :: ::
Db      61 enipqlfcdlvfdkefsfl-tlgekevftlldtklptlshkailrkkrkkrkk 112
```

RESULT 13
AAG25823
ID AAG25823 standard; Protein; 208 AA.
XX
AC AAG25823;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 30043.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 29.3%; Score 374.5; DB 21; Length 208;
Best Local Similarity 39.4%; Pred. No. 7e-26;
Matches 86; Conservative 29; Mismatches 62; Indels 41; Gaps 5;

```
Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
           |||||:|||:|||: : |    |:|:|| :||| ||   | : |:| ||:  |::||||
Db       1 mefwgvevkngkplhldpgldrlvhisqvalgesknnvtepiqlyvtvgsdklligtlsh 60

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGNAAK 120
           |  |||  ::| :: | |||||  |||:| |||    :|    | |:  :::: |    |
Db      61 ekfpqlsteivlernfalshtwkngsvffsgykvdasdp----epedliddqleaa-gfk 115

Qy     121 AVAKPKAKPA-------EVKPAVDDEEDESDSDGMDEDDS---------- 153
           |  |  ||          :||   ||: | |: |  |:|||
Db     116 aapksaakqvnfqlpnedvkakqdddadgseedssddddsensgdeeeekvtaesdseed 175

Qy     154 ----DGEDSEEEEPTPKKPASSKKRANETTPKAPVSAK 187
               | ||   || |||||   |||: |     | |:|
Db     176 dssddeeddsseeetpkkpeepkkrsae-----pnssk 208
```

RESULT 14
AAG09171
ID AAG09171 standard; Protein; 203 AA.
XX
AC AAG09171;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 6998.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.

PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 23.9%; Score 306; DB 21; Length 203;
Best Local Similarity 36.8%; Pred. No. 1e-19;
Matches 86; Conservative 28; Mismatches 80; Indels 40; Gaps 12;

```
Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGEC-KNKKGEFVPLHVKVGN--QNLVLGT 57
           |||||||:| |||  |  ::| ::| || :||:  | || |   ::||:|:     ::|
Db       1 mefwgieikpgkpfkviqkdgfmvhasqvtlgdvekvkkdetfavyvkigddengfmign 60

Qy      58 LSTENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGN 117
           || :  ||   ||    |||:||     ||| :||:|   :     :|| : : |:
Db      61 ls-qkipqfsidlylghefeishn-stssvyligyrt--fdafdeldeeidsdseldeym 116

Qy     118 AAKAVAKPKAKPAEVKPAVDDEEDESDSD--GMDEDD--SDGEDSEEEEPTPKKPASSKK 173
             :  | |:    |: |   :|:|||||  |:|||  || || | | |    | | |
Db     117 eqqlaalpq---neinp---eeddesdsdemgldedddssdeedveaeaplkvappskk- 169

Qy     174 RANETTPKAPVSAKKAKVAVTPQKTDEKKKGGKAANQS-PKSASQVSCG-SCKK 225
                           |    |  ||||    |  |      || ||||
Db     170 --------------------mpngafeiakggkknkssggkkrcpfpcgpsckk 203
```

RESULT 15
AAG17704
ID AAG17704 standard; Protein; 155 AA.
XX
AC AAG17704;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 18826.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.

PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 20.0%; Score 255.5; DB 21; Length 155;
Best Local Similarity 47.3%; Pred. No. 2.6e-15;
Matches 71; Conservative 10; Mismatches 32; Indels 37; Gaps 9;

```
Qy      89 FVGYKTPNIEPQGY-SEEEEEEEEVPA---------GNAAKAV----AKPKAKPAEVKP 134
           |   | ||||   : | ::|:  | |||          |||  ||    :|||||||||||
Db       6 faskkkpnieqddftssddedvpeavpapaptavtangnagaavvkadskpkakpaevkp 65

Qy     135 AVD----DEEDESDSD-----------GM--DEDDSDG---EDS--EEEEPTPKKPAS-S 171
           | :    |||||||| :           ||  ||||||    |||  |||| || ||   :
Db      66 aeekpesdeedesddedeseedddsekgmdvdeddsdddeeedsedeeeeetpnkpepin 125

Qy     172 KKRANETTPKAPVSAKKAKVAVTPQKTDEK 201
           ||| ||:    ||| |||| |  |  | :|
Db     126 kkrpnesvsxtpvsgkkakpaaapastpqk 155
```

Search completed: May 2, 2002, 22:03:52
Job time: 6540 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: May 2, 2002, 20:26:44 ; Search time 60.74 Seconds
(without alignments)
98.523 Million cell updates/sec

Title: US-09-645-337-6
Perfect score: 1279
Sequence: 1 MEFWGIEVKSGKPVTVTPEE.....TFNSGNALESHNKAKHAAAK 245

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	513.5	40.1	311	4	US-09-282-305-16	Sequence 16, Appl
2	499	39.0	302	4	US-09-282-305-14	Sequence 14, Appl
3	487	38.1	305	4	US-09-282-305-12	Sequence 12, Appl
4	446	34.9	285	4	US-09-282-305-18	Sequence 18, Appl
5	224	17.5	412	2	US-08-741-134-2	Sequence 2, Appli
6	166.5	13.0	714	2	US-08-990-114-3	Sequence 3, Appli
7	166.5	13.0	714	4	US-09-241-333-3	Sequence 3, Appli
8	149	11.6	257	4	US-09-538-871-4	Sequence 4, Appli
9	139.5	10.9	411	2	US-08-741-134-6	Sequence 6, Appli
10	136.5	10.7	292	4	US-09-538-871-2	Sequence 2, Appli
11	129.5	10.1	1085	1	US-08-431-080-28	Sequence 28, Appl
12	129.5	10.1	1085	2	US-08-938-534-28	Sequence 28, Appl
13	127	9.9	565	4	US-08-961-083-218	Sequence 218, App
14	126.5	9.9	258	4	US-08-961-083-90	Sequence 90, Appl
15	126.5	9.9	723	1	US-07-814-964-11	Sequence 11, Appl
16	126.5	9.9	723	1	US-08-258-442-11	Sequence 11, Appl
17	126.5	9.9	723	1	US-08-328-809-6	Sequence 6, Appli
18	126.5	9.9	723	5	PCT-US92-11107-11	Sequence 11, Appl
19	121.5	9.5	240	2	US-08-114-555A-8	Sequence 8, Appli
20	121.5	9.5	240	3	US-08-559-397A-14	Sequence 14, Appl
21	121.5	9.5	251	2	US-08-766-738-3	Sequence 3, Appli
22	121	9.5	542	1	US-08-701-380-2	Sequence 2, Appli
23	121	9.5	542	3	US-09-032-365A-13	Sequence 13, Appl
24	117.5	9.2	226	1	US-08-431-080-26	Sequence 26, Appl
25	117.5	9.2	226	2	US-08-938-534-26	Sequence 26, Appl
26	116.5	9.1	739	4	US-09-022-983-2	Sequence 2, Appli
27	115	9.0	764	1	US-08-375-300-4	Sequence 4, Appli
28	115	9.0	764	3	US-09-177-431-4	Sequence 4, Appli
29	115	9.0	764	5	PCT-US95-16930-4	Sequence 4, Appli
30	115	9.0	1089	1	US-08-375-300-2	Sequence 2, Appli
31	115	9.0	1089	3	US-09-177-431-2	Sequence 2, Appli
32	115	9.0	1089	5	PCT-US95-16930-2	Sequence 2, Appli
33	113	8.8	674	3	US-08-893-852A-1	Sequence 1, Appli
34	110.5	8.6	740	4	US-09-022-983-5	Sequence 5, Appli
35	109.5	8.6	160	2	US-08-726-306A-68	Sequence 68, Appl
36	109.5	8.6	251	2	US-08-766-738-1	Sequence 1, Appli
37	109.5	8.6	1187	1	US-08-320-559-28	Sequence 28, Appl
38	109.5	8.6	1187	3	US-08-545-860D-28	Sequence 28, Appl
39	109.5	8.6	1187	5	PCT-US94-04496-28	Sequence 28, Appl
40	109.5	8.6	1210	1	US-08-320-559-26	Sequence 26, Appl
41	109.5	8.6	1210	3	US-08-545-860D-26	Sequence 26, Appl
42	109.5	8.6	1210	5	PCT-US94-04496-26	Sequence 26, Appl
43	109.5	8.6	1706	2	US-08-459-568-2	Sequence 2, Appli
44	109.5	8.6	1706	2	US-08-399-411-2	Sequence 2, Appli
45	109.5	8.6	1706	3	US-08-516-859A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-282-305-16
; Sequence 16, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
; APPLICANT: Baldwin, Donald A.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Crane, Virginia C.
; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
; FILE REFERENCE: 5718-44,
; CURRENT APPLICATION NUMBER: US/09/282,305
; CURRENT FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 60/080,563
; PRIOR FILING DATE: 1998-04-03
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Zea mays
US-09-282-305-16

Query Match 40.1%; Score 513.5; DB 4; Length 311;
Best Local Similarity 42.3%; Pred. No. 8.1e-40;
Matches 127; Conservative 29; Mismatches 81; Indels 63; Gaps 9;

```
Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
           |||||:||| |  |    |  | ::|||||:|||  :|| :   ::||| :: | :||||
Db       1 MEFWGLEVKPGSTVKCEPGHGFILHVSQAALGE--SKKSDSALMYVKVDDKKLAIGTLSI 58

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNI---EPQGYSEEEEEEEEEVPA-- 115
           :  ||:  ||||:|||||||||     ||:| |||       :      |:|||| :|
Db      59 DKYPQIQFDLVFNKEFELSHTSKTTSVFFSGYKVEQPIEGDEMDLDSEDEEEELNIPVIK 118

Qy     116 ----------GNAAKAVAKPKAKPA---EVKPAVD-DEEDESDSDGMDEDDSD------ 154
                    |  ||||   :| :   | |   | |: || :|| ||||||
Db     119 ENGKADGKEEQKNQEKAVAATASKSSLGLEKKSKDDSDDSDEDESDDSDEDDSDDSDEGE 178

Qy     155 ----------------GEDSEEEEPTPKKPASSKKRANETTPKAPVSAKKAKVAVTP-Q 196
                          :| ||| |||||| : |||   |   | |:| ||||||  | |
Db     179 GLSPDEGDDDSSDEDDTSDDDEEETPTPKKPEAGKKRGAENALKTPLSDKKAKVATPPAQ 238

Qy     197 KTDEKK----------KGGKAAN------QSPKSASQVSCGSCKKTFNSGNALESHNKA 239
           ||  ||           ||   ||      :||||   | | || ||||| ||::|:||
Db     239 KTGGKKGATHVATPHPAKGKTPANNDKSTEKSPKSGGSVPCKSCSKTFNSEMALQAHSKA 298
```

RESULT 2
US-09-282-305-14

```
; Sequence 14, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
;  APPLICANT:  Baldwin, Donald A.
;  APPLICANT:  Briggs, Steven P.
;  APPLICANT:  Crane, Virginia C.
;  TITLE OF INVENTION:  Maize Histone Deacetylases And Their Uses
;  FILE REFERENCE:  5718-44,
;  CURRENT APPLICATION NUMBER: US/09/282,305
;  CURRENT FILING DATE:  1999-03-31
;  PRIOR APPLICATION NUMBER:  60/080,563
;  PRIOR FILING DATE:  1998-04-03
;  NUMBER OF SEQ ID NOS:  18
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 302
;   TYPE: PRT
;   ORGANISM: Zea mays
US-09-282-305-14
```

```
  Query Match             39.0%;  Score 499;  DB 4;  Length 302;
  Best Local Similarity   40.5%;  Pred. No. 1.7e-38;
  Matches  124;  Conservative   36;  Mismatches   74;  Indels   72;  Gaps   11;

Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
             |||||:|||  |   |    |  | ::|:|||:|||  :|| :   ::|||: :| | :||||
Db       1 MEFWGLEVKPGSTVKCEPGYGFVLHLSQAALGE--SKKSDNALMYVKIDDQKLAIGTLSV 58

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSE-------EEEEEEEEV 113
           :  | :  ||:||||||||||    ||:| |||       | : |        |:|:||  |
Db      59 DKNPHIQFDLIFDKEFELSHTSKTTSVFFTGYKV----EQPFEEDEMDLDSEDEDEELNV 114

Qy     114 PA----GNA--------AKAVAKP-KAKPAEVKPAVDDEEDESDSDGMDEDDSD------ 154
           |     | |          |||| | |: |    |   ||: || ::|  |||::|
Db     115 PVVKENGKADGKKQKSQEKAVAAPSKSSPDSKKSKDDDDSDEDETDDSDEDETDDSDEGL 174

Qy     155 --------------GEDSEEEEPTPKKPASSKKRANETTP-KAPVSAKKAKVAVTP--Q 196
                         :| ||: ||||||   |||| |::  | |:| |||||| ||  |
Db     175 SSEEGDDDSSDEDDTSDDEEEDTPTPKKPEVGKKRAAESSVLKTPLSDKKAKVA-TPSSQ 233

Qy     197 KTDEKK--------------------KGGKAANQSPKSASQVSCGSCKKTFNSGNALES 235
           ||  ||                    |  |:   :|||   | | || |:| |  ||::
Db     234 KTGGKKGAAVHVATPHPAKGKTIVNNDKSVKSPKSAPKSGGSVPCKSCSKSFISETALQA 293

Qy     236 HNKAKH 241
           |:|||
Db     294 HSKAKH 299
```

```
RESULT   3
US-09-282-305-12
; Sequence 12, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
;  APPLICANT:  Baldwin, Donald A.
;  APPLICANT:  Briggs, Steven P.
;  APPLICANT:  Crane, Virginia C.
;  TITLE OF INVENTION:  Maize Histone Deacetylases And Their Uses
;  FILE REFERENCE:  5718-44,
;  CURRENT APPLICATION NUMBER: US/09/282,305
;  CURRENT FILING DATE:  1999-03-31
;  PRIOR APPLICATION NUMBER:  60/080,563
;  PRIOR FILING DATE:  1998-04-03
;  NUMBER OF SEQ ID NOS:  18
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 12
;   LENGTH: 305
;   TYPE: PRT
;   ORGANISM: Zea mays
US-09-282-305-12
```

```
  Query Match             38.1%;  Score 487;  DB 4;  Length 305;
  Best Local Similarity   41.2%;  Pred. No. 2.2e-37;
  Matches  127;  Conservative   37;  Mismatches   74;  Indels   70;  Gaps   13;

Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
             |||||:|||  |   |    |  | ::|:|||:|||  :|| :   ::|||: :| | :||||
Db       1 MEFWGLEVKPGSTVKCEPGYGFVLHLSQAALGE--SKKSDNALMYVKIDDQKLAIGTLSV 58

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSE-------EEEEEEEEV 113
           :  | :  ||:||||||||||    ||:| |||       | : |        |:|:||  |
Db      59 DKNPHIQFDLIFDKEFELSHTSKTTSVFFTGYKV----EQPFEEDEMDLDSEDEDEELNV 114

Qy     114 PA----GNA--------AKAVAKP-KAKPAEVKPAVDDEEDESDSDGMDEDDSD------ 154
           |     | |          |||| | |: |    |   ||: || ::|  |||::|
Db     115 PVVKENGKADGKKQKSQEKAVAAPSKSSPDSKKSKDDDDSDEDETDDSDEDETDDSDEGL 174

Qy     155 --------------GEDSEEEEPTPKKPASSKKRANETTP-KAPVSAKKAKVAVTP--Q 196
                         :| ||: ||||||   |||| |::  | |:| |||||| ||  |
Db     175 SPEEGDDDSSDEDDTSDDEEEDTPTPKKPEVGKKRAAESSVLKTPLSDKKAKVA-TPSSQ 233

Qy     197 KTDEKK------------KGGKAAN-----QSPKSA--SQVSCGSCKKTFNSGNALESHN 237
           ||  ||              ||    |      :|||||  | | | || |:| |  ||::|:
Db     234 KTGGKKGAAVHVATPHPAKGKTIVNNDKSVKSPKSAPKSGVPCKSCSKSFISETALQAHS 293

Qy     238 KAKHAAAK 245
           ||   |::
Db     294 KANMGASE 301
```

```
RESULT   4
US-09-282-305-18
; Sequence 18, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
;  APPLICANT:  Baldwin, Donald A.
;  APPLICANT:  Briggs, Steven P.
;  APPLICANT:  Crane, Virginia C.
;  TITLE OF INVENTION:  Maize Histone Deacetylases And Their Uses
;  FILE REFERENCE:  5718-44,
;  CURRENT APPLICATION NUMBER: US/09/282,305
;  CURRENT FILING DATE:  1999-03-31
;  PRIOR APPLICATION NUMBER:  60/080,563
;  PRIOR FILING DATE:  1998-04-03
;  NUMBER OF SEQ ID NOS:  18
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 18
;   LENGTH: 285
;   TYPE: PRT
;   ORGANISM: Zea mays
US-09-282-305-18
```

```
  Query Match             34.9%;  Score 446;  DB 4;  Length 285;
  Best Local Similarity   41.9%;  Pred. No. 1.2e-33;
  Matches  122;  Conservative   31;  Mismatches   84;  Indels   54;  Gaps   12;

Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
             |||||  |||  |  |:    :|::||:|||:|||  |  |   |   : ||: :: |||||||
Db       1 MEFWGEEVKPGATVSCKVGDGLVIHLSQAALGEPK-KASENAIVSVKIDDKKLVLGTLSV 59

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTP----NIEPQGYSEEEEEEEEEVPAG 116
           |  ||: |||||||:|||||      ||:| |||:|      :     | :|| | : :|
Db      60 EKHPQISCDLVFDKDFELSHNSKTASVFFCGYKSPVPLFESDSGEDSSDEEVEPDLIPMQ 119

Qy     117 N------AAKAVAK---PKAKPAEVKPAVDDEEDESDSDGMDEDDSDGED---------- 157
           |        ||   |      |   |      ||:  :|||:  :||:|  ||
Db     120 NNEIKISTAKVPVKVGIQNADEDETSSGDDDDFTDSDSEMSEEDESSDEDEVSSDTDTSD 179

Qy     158 ---SEEEEPTPKKP--ASSKKRANET-TPKAPVSAKKAKVAVTPQKTDEK-------KKG 204
              |||: |||||      |||| |  ||    | ||||   : ||| :|       |:
Db     180 DSGSEEQTPTPKKTEVVVGKKRAIEAETP----SGKKAKSEQSAQKTGDKVSTSHPAKQS 235

Qy     205 GK-----------AANQSPKSASQVSCGSCKKTFNSGNALESHNKA-KHAA 243
```

```
Db    236 SKTPADKSTKTPTADKKSPKSGSH-ACKSCSKSFGSASALESHQKAKKHEA 285


RESULT   5
US-08-741-134-2
; Sequence 2, Application US/08741134
; Patent No. 5861498
;  GENERAL INFORMATION:
;    APPLICANT:  Litwack, Gerald
;    APPLICANT:  Alnemri, Emad S.
;    APPLICANT:  Fernandes-Alnemri, Teresa
;    TITLE OF INVENTION:  IMMUNOPHILIN FKBP46 AND COMPOSITIONS FOR MAKING
;    TITLE OF INVENTION:  AND
;    TITLE OF INVENTION:  METHODS OF USING THE SAME
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz & No. 5861498ris
;      STREET:  One Liberty Place - 46th floor
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;.     ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 3.11
;      SOFTWARE:  WordPerfect for Windows 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/741,134
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/007,163
;      FILING DATE:  01-NOV-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DeLuca, Mark
;      REGISTRATION NUMBER:  33,229
;      REFERENCE/DOCKET NUMBER:  TJU-2090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH: 412 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-741-134-2


  Query Match           17.5%;  Score 224;  DB 2;  Length 412;
  Best Local Similarity  30.1%;  Pred. No. 6.6e-13;
  Matches   82;  Conservative   31;  Mismatches   91;  Indels   68;  Gaps   10;

Qy      3 FWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLSTEN 62
          |||: ::  |  |   |:    |:|||:: :      :   : | |  :| :: ||
Db      2 FWGLIMEPNKRYTQVVEKP--FHISQAAM-DISTGDNDPCQVMVVVDGKNFLVCTLQKGK 58

Qy     63 IPQLFCDLVFDKEFELSH-TWGKGSVYFVGYKTPNIEPQGYSEEEEEEEEE-------VP 114
          | |:  || |     :|  | || :|:  ||  |  |     ||| |||||        ||
Db     59 IIQVPLDLYFKSGDSVSFLTNGKCNVHLTGYLDPEFEEDLEDEEEAEEEEEEEEAPPLVP 118

Qy    115 AGNAAK------AVAKPKAKP-----------AEVKPAVDDEE----------------- 140
          | |  |      | |  ||||           ||     |||:
Db    119 AKNKRKLENANDATANKKAKPDKKAGKNSAPAAESDSDDDDEDQLQKFLDGEDIDTDEND 178

Qy    141 -----------DESDSDGMDED----DSDGEDSEEEEPTPKK----PASSKKRANETTPK 181
                      |:|| :  |||    : | || ||||  |||    ||: :    :|: :
Db    179 ESFKMNTSAEGDDSDEEDDDEDEEDEEDDDEDDEEEEEAPKKKKKQPAAEQDSTLDTSKE 238

Qy    182 APVSAKKAKVAVTPQKTDEKKKGGKAANQSPK 213
```

```
          :   :| :|     ||   |||  : | | |:
Db    239 SVDMSKLSK----SQKRRLKKKLQQQAKQQPQ 266


RESULT   6
US-08-990-114-3
; Sequence 3, Application US/08990114
; Patent No. 5932475
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  HUMAN NUCLEOLIN-LIKE PROTEIN
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/990,114
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0451 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  714 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  128842
US-08-990-114-3


  Query Match           13.0%;  Score 166.5;  DB 2;  Length 714;
  Best Local Similarity  33.8%;  Pred. No. 2.8e-07;
  Matches   48;  Conservative   20;  Mismatches   55;  Indels   19;  Gaps    4;

Qy     73 DKEFELSHTWGK-GSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGNAAKAV----AKPKA 127
          : |||      || | |       : : : :   |||||:|||        :|:     || |
Db    166 EDEFEPPVVKGKQGKVAAAAPASEDEDEEEDEEEEEEDEEEEDDSEEEEAMEITPAKGKK 225

Qy    128 KPAEVKPAV----------DDEEDESDSDGMDEDDSDGEDSEEEEPTPKKPASSKKRANE 177
           ||:| |             |:|||| : :  :|:: : |: |||| | |||  |::
Db    226 APAKVVPVKAKNVAEEDDDDEEEDEDEEEDEEEEEDEEEEEEEEEEPVKPAPGKRKKEM 285

Qy    178 TTPKAPVSAKKAKV----AVTP 195
          |  |    ||| ||    : ||
Db    286 TKQKEVPEAKKQKVEGSESTTP 307


RESULT   7
US-09-241-333-3
```

```
; Sequence 3, Application US/09241333
; Patent No. 6313266
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  HUMAN NUCLEOLIN-LIKE PROTEIN
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/241,333
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/990,114
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0451 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  714 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  128842
US-09-241-333-3


  Query Match             13.0%;  Score 166.5;  DB 4;  Length 714;
  Best Local Similarity   33.8%;  Pred. No. 2.8e-07;
  Matches   48;  Conservative   20;  Mismatches   55;  Indels   19;  Gaps    4;

Qy      73 DKEFELSHTWGK-GSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGNAAKAV----AKPKA 127
           : |||     || | |      : : : :   |||||:|||       :|:    || |
Db     166 EDEFEPPVVKGKQGKVAAAAPASEDEDEEEDEEEEEEDEEEEDDSEEEEAMEITPAKGKK 225

Qy     128 KPAEVKPAV----------DDEEDESDSDGMDEDDSDGEDSEEEEPTPKKPASSKKRANE 177
             ||:| |            |:|||| : :  :|:: : |: ||||  | |||  |::
Db     226 APAKVVPVKAKNVAEEDDDDEEEDEDEEEDEEEEEDEEEEEEEEEEEPVKPAPGKRKKEM 285

Qy     178 TTPKAPVSAKKAKV----AVTP 195
           |  |    ||| ||    : ||
Db     286 TKQKEVPEAKKQKVEGSESTTP 307


RESULT   8
US-09-538-871-4
; Sequence 4, Application US/09538871
; Patent No. 6358718
; GENERAL INFORMATION:
;  APPLICANT: Olson, Mark
;  APPLICANT: Szebeni, Atilla
;  TITLE OF INVENTION: METHODS FOR STABILIZATION AND
;  TITLE OF INVENTION: RENATURATION OF PROTEINS USING NUCLEOLAR PROTEINS B23
;  FILE REFERENCE: 13121.0002U2
;  CURRENT APPLICATION NUMBER: US/09/538,871
;  CURRENT FILING DATE: 2000-03-30
;  EARLIER APPLICATION NUMBER: 60/126,910
;  EARLIER FILING DATE: 1999-03-30
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 257
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of artificial sequence:/No. 6358718e -
;   OTHER INFORMATION: synthetic construct
US-09-538-871-4


  Query Match             11.6%;  Score 149;  DB 4;  Length 257;
  Best Local Similarity   27.9%;  Pred. No. 3e-06;
  Matches   39;  Conservative   26;  Mismatches   69;  Indels    6;  Gaps    2;

Qy      83 GKGSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGNAAKAVAKPKAKPAEVKPAVDDEEDE 142
           | | |:  |     :|    ||:|:|||: ::   :  ::      |  : |  :|:::||
Db     105 GSGPVHISGQHLVAVEEDAESEDEDEEDVKLLGMSGKRSAPGGGNKVPQKKVKLDEDDDE 164

Qy     143 SDSDGMDEDDSDGEDSEEEEPTPKKPASSKKRANET----TPKAPVSAKKAKVAVTPQKT 198
           | |  |::| | :| :|||   | |   ||   :|      |:  : |  | :     |
Db     165 DDEDDEDDEDDDDDDFDEEETEEKVPV--KKSVRDTPAKNAQKSNQNGKDLKPSTPRSKG 222

Qy     199 DEKKKGGKAANQSPKSASQV 218
           |  |  :   ::||  | |
Db     223 QESFKKQEKTPKTPKGPSSV 242


RESULT   9
US-08-741-134-6
; Sequence 6, Application US/08741134
; Patent No. 5861498
;  GENERAL INFORMATION:
;    APPLICANT:  Litwack, Gerald
;    APPLICANT:  Alnemri, Emad S.
;    APPLICANT:  Fernandes-Alnemri, Teresa
;    TITLE OF INVENTION:  IMMUNOPHILIN FKBP46 AND COMPOSITIONS FOR MAKING
;    TITLE OF INVENTION:  AND
;    TITLE OF INVENTION:  METHODS OF USING THE SAME
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz & No. 5861498ris
;      STREET:  One Liberty Place - 46th floor
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 3.11
;      SOFTWARE:  WordPerfect for Windows 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/741,134
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/007,163
;      FILING DATE:  01-NOV-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DeLuca, Mark
;      REGISTRATION NUMBER:  33,229
;      REFERENCE/DOCKET NUMBER:  TJU-2090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
```

```
;      TELEFAX:  215-568-3439
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;      LENGTH: 411 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-741-134-6


  Query Match            10.9%;  Score 139.5;  DB 2;  Length 411;
  Best Local Similarity  27.1%;  Pred. No. 4.2e-05;
  Matches   60;  Conservative   22;  Mismatches   98;  Indels   41;  Gaps    8;

Qy      51 QNLVLGTLSTENIPQLFCDLVFDKEFELSH-TWGKGSVYFVG--YKTPNIEPQGYSEEEE 107
           |  || ||| |   |   ||    | |:     |  ::   |   | |   | |   |:|
Db     118 QESVLLTLSPEAQYQQSLDLTITPEEEVQFIVTGSYAISLSGNYVKHPFDTPMGVEGEDE 177

Qy     108 EEEEEVPAGNAAKAVAKPKAKPAEVKPAVDDEEDE----SDSDGMDEDDSDGEDSEEEE- 162
           :|: ::                 :    :||||:|     :    ||:|:|||: :|||
Db     178 DEDADIYDSEDYDLTPDEDEIIGDDMDDLDDEEEEEVRIEEVQEEDEEDNDGEEEQEEEE 237

Qy     163 ---------PTPKKPASSKKRANETTPKAPVSAKKAKVAVT------PQKTDEKKKGGKA 207
                   | |||    ||| :|   :   : |  ||          | |   ||:  |
Db     238 EEEQKEEVKPEPKKSKKEKKRKHEEKEEEKKAKKVKKVEFKKDLEEGPTKPKSKKEQDK- 296

Qy     208 ANQSPKSASQVSCGSCKKTFNSGNALESHN-----KAKHAA 243
               |||           |    |  :|         :||  |
Db     297 --HKPKS----------KVLEGGIVIEDRTIGDGPQAKRGA 325


RESULT  10
US-09-538-871-2
; Sequence 2, Application US/09538871
; Patent No. 6358718
; GENERAL INFORMATION:
;  APPLICANT: Olson, Mark
;  APPLICANT: Szebeni, Atilla
;  TITLE OF INVENTION: METHODS FOR STABILIZATION AND
;  TITLE OF INVENTION: RENATURATION OF PROTEINS USING NUCLEOLAR PROTEINS B23
;  FILE REFERENCE: 13121.0002U2
;  CURRENT APPLICATION NUMBER: US/09/538,871
;  CURRENT FILING DATE: 2000-03-30
;  EARLIER APPLICATION NUMBER: 60/126,910
;  EARLIER FILING DATE: 1999-03-30
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 292
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of artificial sequence:/No. 6358718e -
;   OTHER INFORMATION: synthetic construct
US-09-538-871-2


  Query Match            10.7%;  Score 136.5;  DB 4;  Length 292;
  Best Local Similarity  25.4%;  Pred. No. 5e-05;
  Matches   52;  Conservative   34;  Mismatches   90;  Indels   29;  Gaps    8;

Qy      24 IHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLSTENIPQLFCDLVFDKEFELSH--- 80
           :|: :|   |  | :|   |: |        | ||     | :   |     ||::
Db      57 LHIVEA---EAMNYEGS--PIKV-------TLATLKMSVQPTVSVHL---GGFEITPPVV 101

Qy      81 ---TWGKGSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGNAAKAVAKPKAKPAEVKPAVD 137
              | |  :   |     :|    ||:|:||: ::   :  ::      | : |  :|
Db     102 LRLKCGSGPI--SGQHLVAVEEDAESEDEDEEDVKLLGMSGKRSAPGGGNKVPQKKVKLD 159

Qy     138 DEEDESDSDGMDEDDSDGEDSEEEEPTPKKPASSKKRANET----TPKAPVSAKKAKVAV 193
           :::|| | |  |::| | :| :|||   | |   ||   :|      |:  : |  | :
Db     160 EDDDEDDEDDEDDEDDDDDDFDEEETEEKVPV--KKSVRDTPAKNAQKSNQNGKDLKPST 217

Qy     194 TPQKTDEKKKGGKAANQSPKSASQV 218
              |  |  |  :   ::||  | |
Db     218 PRSKGQESFKKQEKTPKTPKGPSSV 242


RESULT  11
US-08-431-080-28
; Sequence 28, Application US/08431080
; Patent No. 5698686
;  GENERAL INFORMATION:
;    APPLICANT:  Gottschling, Daniel E.
;    APPLICANT:  Singer, Miriam S.
;    TITLE OF INVENTION:  Telomerase Compositions and Methods
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  TEXAS
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS/ASCII
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/431,080
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  SN 08/326,781
;      FILING DATE:  October 20, 1994
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parker, David L.
;      REGISTRATION NUMBER:  32,165
;      REFERENCE/DOCKET NUMBER:  ARCD:155/PAR
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (713) 789-2679
;      TELEX:  79-0924
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1085 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-431-080-28


  Query Match            10.1%;  Score 129.5;  DB 1;  Length 1085;
  Best Local Similarity  23.5%;  Pred. No. 0.0013;
  Matches   62;  Conservative   29;  Mismatches   96;  Indels   77;  Gaps    9;

Qy      15 TVTPEEGILIHVSQASLGEC----KNKKGEFVPLHVKVGNQNLVLGTLSTENIPQLFCDL 70
           |  |     ||: |::|| :     |||     |  :|   :|:         |  |
Db      47 TTRPRRFSLIYSSESSLSDVSDSDKNKSTN--PHKIKRKAKNI-------SNNSQGKKSK 97

Qy      71 VFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGNAAKAVAKPKAKPA 130
           :  :: :     : | |     |   | :   || ||||||:
Db      98 LIQRQIDNDDEGTESSDYQA--VTDGEESENEEEESEEEEED----------------- 137

Qy     131 EVKPAVDDEEDESDSDGMDEDDSDGEDSEEEEPTP--KKPASSKKRA----------NE 177
                 |||:|: | |    : ||| | | ::|     |  |  ||||           |
Db     138 ------DDEDDDDDDDDGSDSDSDSETSSDDENIDFVKLTAQRKKRAMKALSAMNTNSNT 191

Qy     178 TTPKAPVSAKKAKVAVTPQKTDEKKKGGKAANQSPKSASQVSCGSCKKTFNS-------- 229
                 | |   | ::|:| :|::  |   :  |   |    | ||  |
Db     192 LYSSRENSNKNKSVKLSPKKENEEEQ--KEEKEKEKEEQQKQQESNKKEVNGSGTTTTQQ 249

Qy     230 --------------GNALESHNK 238
                          ||  | :|:
```

```
Db    250 ALSFKFKKEDDGISFGNGNEGYNE 273


RESULT 12
US-08-938-534-28
; Sequence 28, Application US/08938534
; Patent No. 5916752
;  GENERAL INFORMATION:
;    APPLICANT:  Gottschling, Daniel E.
;    APPLICANT:  Singer, Miriam S.
;    TITLE OF INVENTION:  Telomerase Compositions and Methods
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  TEXAS
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS/ASCII
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/938,534
;      FILING DATE:  26-SEP-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/431,080
;      FILING DATE:
;      APPLICATION NUMBER:  SN 08/326,781
;      FILING DATE:  October 20, 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parker, David L.
;      REGISTRATION NUMBER:  32,165
;      REFERENCE/DOCKET NUMBER:  ARCD:155/PAR
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (713) 789-2679
;      TELEX:  79-0924
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1085 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-938-534-28


  Query Match             10.1%;  Score 129.5;  DB 2;  Length 1085;
  Best Local Similarity   23.5%;  Pred. No. 0.0013;
  Matches   62;  Conservative   29;  Mismatches   96;  Indels   77;  Gaps    9;

Qy      15 TVTPEEGILIHVSQASLGEC----KNKKGEFVPLHVKVGNQNLVLGTLSTENIPQLFCDL 70
           |  |    ||: |::|| :     |||     |  :|   :|:         |  |
Db      47 TTRPRRFSLIYSSESSLSDVSDSDKNKSTN--PHKIKRKAKNI-------SNNSQGKKSK 97

Qy      71 VFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGNAAKAVAKPKAKPA 130
             :  :: :      : | |     |   | :   || |||||:
Db      98 LIQRQIDNDDEGTESSDYQA--VTDGEESENEEEESEEEEED----------------- 137

Qy     131 EVKPAVDDEEDESDSDGMDEDDSDGEDSEEEEPTP--KKPASSKKRA----------NE 177
                ||:|: | |   : |||  | | ::|     |  |  ||||           |
Db     138 ------DDEDDDDDDDDGSDSDSDSETSSDDENIDFVKLTAQRKKRAMKALSAMNTNSNT 191

Qy     178 TTPKAPVSAKKAKVAVTPQKTDEKKKGGKAANQSPKSASQVSCGSCKKTFNS-------- 229
               | |   | ::|:| :|:::  |   :  |   |    | ||  |
Db     192 LYSSRENSNKNKSVKLSPKKENEEEQ--KEEKEKEKEEQQKQQESNKKEVNGSGTTTTQQ 249

Qy     230 --------------GNALESHNK 238
                          ||  | :|:
Db     250 ALSFKFKKEDDGISFGNGNEGYNE 273


RESULT 13
US-08-961-083-218
; Sequence 218, Application US/08961083
; Patent No. 6159469
;  GENERAL INFORMATION:
;    APPLICANT:  Choi et. al.
;    TITLE OF INVENTION:  Streptococcus pneumoniae Antigens and Vaccines
;    NUMBER OF SEQUENCES:  452
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/961,083
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB340P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  218:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  565 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-961-083-218


  Query Match             9.9%;  Score 127;  DB 4;  Length 565;
  Best Local Similarity   30.3%;  Pred. No. 0.00092;
  Matches   43;  Conservative   18;  Mismatches   61;  Indels   20;  Gaps    5;

Qy      90 VGYKTPNI---EPQGYSEEEEEEEEEVPAGNAAKAVAKPKAKP-AEVKPAVD-------D 138
           |  :||     : |  :| :  | || |   | |    || :| :||||  |
Db       2 VEVETPQSITNQEQARTENQVVETEEAPKEEAPKTEESPKEEPKSEVKPTDDTLPKVEEG 61

Qy     139 EEDESDSDGMDEDDSDGEDSEEEEPTPK-------KPASSKKRANETTPKAPVSAKKAKV 191
              :|| ::   ::|   : |   ||:   |       |||   |       | ||   :||
Db      62 KEDSAEPAPVEEVGGEVESKPEEKVAVKPESQPSDKPAEESKVEQAGEPVAPREDEKA-- 119

Qy     192 AVTPQKTDEKKKGGKAANQSPK 213
               | |:|  |  :  ||  ::||
Db     120 PVEPEKQPEAPEEEKAVEETPK 141


RESULT 14
US-08-961-083-90
; Sequence 90, Application US/08961083
; Patent No. 6159469
;  GENERAL INFORMATION:
;    APPLICANT:  Choi et. al.
;    TITLE OF INVENTION:  Streptococcus pneumoniae Antigens and Vaccines
;    NUMBER OF SEQUENCES:  452
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
```

```
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/961,083
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB340P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  90:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  258 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-961-083-90


  Query Match             9.9%;  Score 126.5;  DB 4;  Length 258;
  Best Local Similarity   30.5%;  Pred. No. 0.00036;
  Matches   40;  Conservative   17;  Mismatches   57;  Indels   17;  Gaps    4;

Qy      98 EPQGYSEEEEEEEEEVPAGNAAKAVAKPKAKP-AEVKPAVD-------DEEDESDSDGMD 149
           :  :| :  | || |   | |    || :| :|||  |         :|, ::   ::
Db       9 QEQARTENQVVETEEAPKEEAPKTEESPKEEPKSEVKPTDDTLPKVEEGKEDSAEPAPVE 68

Qy     150 EDDSDGEDSEEEEPTPK-------KPASSKKRANETTPKAPVSAKKAKVAVTPQKTDEKK 202
           |   : |   ||:   |       |||   |      | ||   :||   | |:|  |
Db      69 EVGGEVESKPEEKVAVKPESQPSDKPAEESKVEQAGEPVAPREDEKA--PVEPEKQPEAP 126

Qy     203 KGGKAANQSPK 213
           :  ||  ::||
Db     127 EEEKAVEETPK 137


RESULT  15
US-07-814-964-11
; Sequence 11, Application US/07814964
; Patent No. 5359047
;  GENERAL INFORMATION:
;    APPLICANT:  Donahue, Brian A.
;    APPLICANT:  Toney, Jeffrey H.
;    APPLICANT:  Bruhn, Suzanne L.
;    APPLICANT:  Pil, Pieter M.
;    APPLICANT:  Brown, Steven
;    APPLICANT:  Kellett, Patti
;    APPLICANT:  Essigmann, John M.
;    APPLICANT:  Lippard, Stephen J.
;    TITLE OF INVENTION:  DNA Structure Specific Recognition
;    TITLE OF INVENTION:  Protein and Uses Therefor
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  2 Militia Drive
;      CITY:  Lexington
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/814,964
;      FILING DATE:  19911226
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/539,906
;      FILING DATE:  18-JUN-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  MIT-4787AAA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  11
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  723 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Drosophila melanogaster
;    IMMEDIATE SOURCE:
;      CLONE:  Drosophila SSRP (predicted)
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  458..507
;      OTHER INFORMATION:  /label= Acidic
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  518..547
;      OTHER INFORMATION:  /label= Basic 1
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  547..620
;      OTHER INFORMATION:  /label= HMG-box
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  632..649
;      OTHER INFORMATION:  /label= Basic II
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  657..723
;      OTHER INFORMATION:  /label= Mixed Charge
US-07-814-964-11


  Query Match             9.9%;  Score 126.5;  DB 1;  Length 723;
  Best Local Similarity   26.4%;  Pred. No. 0.0014;
  Matches   58;  Conservative   36;  Mismatches   85;  Indels   41;  Gaps   12;

Qy      54 VLGTLSTENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNI-------EPQGY---- 102
           :  ::  |   :|| | :  |:  :|:  ||      |||  :        ||  |
Db     402 IFSSIEKEEYAKLF-DYITQKKLHVSN-MGKDK---SGYKDVDFGDSDNENEPDAYLARL 456

Qy     103 SEEEEEEEEEVPAGNAAKAVAKPKAKPAEVKPAVDDEED---ESDSDGMDED------DS 153
             |  |:||:   |:: :       || | :  | :| |   |||||  | |      ||
Db     457 KAEAREKEEDDDDGDSDEESTDEDFKPNENESDVAEEYDSNVESDSDD-DSDASGGGGDS 515

Qy     154 DGEDSEEEEPTPKKPASSKKRANETTPKAPVSAKKAKVAVTPQKT--------DEKKKGG 205
           ||   ::|: : ||    ||   :   | |  :|| | :  |::         :: ::
Db     516 DGAKKKKEKKSEKKEKKEKKHKEKERTKKP--SKKKKDSGKPKRATTAFMLWLNDTRESI 573

Qy     206 KAANQSPKSASQVSCGSCKKTFNSGNALESHNKAKHAAAK 245
           |  |   |  ::::    ||       |:  :| : ||||
Db     574 KRENPGIK-VTEIA----KKGGEMWKELKDKSKWEDAAAK 608
```

Search completed: May 2, 2002, 22:08:41
Job time: 6117 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: May 2, 2002, 20:16:12 ; Search time 205.36 Seconds
(without alignments)
114.637 Million cell updates/sec

Title: US-09-645-337-6
Perfect score: 1279
Sequence: 1 MEFWGIEVKSGKPVTVTPEE.....TFNSGNALESHNKAKHAAAK 245

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	594	46.4	305	2	T52287	probable histone d
2	498.5	39.0	296	2	T48401	histone deacetylas
3	482	37.7	307	2	T04141	histone deacetylas
4	224	17.5	412	2	A55320	immunophilin FKBP4
5	188.5	14.7	357	2	JC4090	FK506-binding 39k
6	170.5	13.3	299	2	A29681	nucleophosmin - Af
7	170	13.3	694	1	DNCHNL	nucleolin - chicke
8	169	13.2	294	1	DNCHFM	nucleophosmin - ch
9	166.5	13.0	713	2	A27441	nucleolin - Chines
10	165.5	12.9	990	2	I51618	nucleolar phosphop
11	164	12.8	487	2	T10215	hypothetical prote
12	157	12.3	1110	2	I51116	NF-180 - sea lampr
13	155	12.1	294	1	A32915	nucleophosmin - hu
14	153.5	12.0	635	2	T09648	nucleolin homolog
15	153.5	12.0	707	1	DNMS	nucleolin - mouse
16	153	12.0	707	2	A35804	nucleolin - human
17	151.5	11.8	712	2	JH0148	nucleolin - rat
18	150.5	11.8	281	2	JC4295	heat-shock protein
19	149.5	11.7	644	2	S55395	neurofilament prot
20	149	11.6	257	2	A34168	nucleolar phosphop
21	149	11.6	292	2	A28939	nucleophosmin - ra
22	148	11.6	296	2	A41730	nucleophosmin NO38
23	146.5	11.5	798	2	T33022	hypothetical prote
24	145.5	11.4	734	2	B42680	nucleolus-cytoplas
25	144	11.3	292	2	I52858	nucleolar protein
26	140.5	11.0	649	2	D85135	hypothetical prote
27	140.5	11.0	798	2	I50479	neurofilament medi
28	140.5	11.0	1881	2	H95076	zinc metalloprotei
29	139.5	10.9	411	2	S48647	peptidylprolyl iso
30	138.5	10.8	971	2	T19431	hypothetical prote
31	138	10.8	699	2	I38073	nucleolar phosphop
32	138	10.8	705	2	S32644	nucleolin - Africa
33	137	10.7	590	2	S63193	hypothetical prote
34	136.5	10.7	292	2	A36089	nucleolar phosphop
35	136.5	10.7	611	2	T06458	nucleolin homolog
36	136.5	10.7	723	2	A48217	single-strand DNA/
37	136	10.6	651	2	S18874	nucleolin - Africa
38	135.5	10.6	557	2	A96527	probable nuM1 prot
39	134.5	10.5	280	2	A61047	ectodermal (ect) -
40	134.5	10.5	325	2	T18283	hypothetical prote
41	134	10.5	1320	2	JC5630	TCOF1 protein - mo
42	133.5	10.4	1876	2	E97944	zinc metalloprotei
43	133	10.4	532	2	T06029	hypothetical prote
44	132.5	10.4	971	2	T24866	hypothetical prote
45	132	10.3	374	2	C88734	protein F32E10.6 [

ALIGNMENTS

RESULT 1
T52287
probable histone deacetylase (EC 3.5.1.-) [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
C;Accession: T52287
R;Dangl, M.; Haas, H.; Loidl, P.
submitted to the EMBL Data Library, January 1998
A;Description: Arabidopsis thaliana complete cDNA-sequence homologous to Zea mays HD2
A;Reference number: Z26016
A;Accession: T52287
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-305 <DAN>
A;Cross-references: EMBL:AF044914; PIDN:AAC02539.1
A;Experimental source: cultivar Columbia
C;Genetics:
A;Gene: HD2
C;Keywords: hydrolase

Query Match 46.4%; Score 594; DB 2; Length 305;
Best Local Similarity 50.7%; Pred. No. 6.5e-31;
Matches 151; Conservative 26; Mismatches 63; Indels 58; Gaps 15;

```
Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVIGTLST 60
           |||||: |       |||||  |:|:|||||  :|  | || | | | ||   ||:||||
Db       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASL-DCTVKSGESVVLSVTVGGAKLVIGTLSQ 59

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGY-SEEEEEEEEVPA---- 115
           :  ||:  |||||||||||:  | :|:|:|||:||||   : | ::|:  | |||
Db      60 DKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPT 119

Qy     116 -----GNAAKAV----AKPKAKPAEVKPAVD----DEEDESDSD----------GM--D 149
                 |||  ||     ||||||||||| :    ||||||| :           ||  |
Db     120 AVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMDVD 179

Qy     150 EDDSDG---EDS--EEEEPTPKKPAS-SKKRANETTPKAPVSAKKAKVAV----TPQKTD 199
           !!!!!    |||  |||| |||||   :||| ||:  | ||| |||| |     |||||:
Db     180 EDDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTE 239

Qy     200 EK----------KKGGKA---ANQSPKSASQVSCG-SCKKTFNSGNAL-ESHNKAKH 241
           :|          |||||:   |||||||  | | | : || ||||     |:||  :
Db     240 KKKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGGNNNKKPFNSGKQFGGSNNKGSN 297
```

RESULT 2
T48401
histone deacetylase-like protein - Arabidopsis thaliana
N;Alternate names: protein F17C15.160
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000

```
C;Accession: T48401
R;Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.;
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24492
A;Accession: T48401
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-296 <BEV>
A;Cross-references: EMBL:AL162506
A;Experimental source: cultivar Columbia; BAC clone F17C15
C;Genetics:
A;Map position: 5
A;Introns: 5/1; 28/3; 93/3; 116/3; 164/3; 195/3; 227/1
A;Note: F17C15.160


  Query Match             39.0%;  Score 498.5;  DB 2;  Length 296;
  Best Local Similarity   39.5%;  Pred. No. 7.4e-25;
  Matches  121;  Conservative   31;  Mismatches   81;  Indels   73;  Gaps    9;

Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
           |||||:|||:|||: : |    |:|:|| :||| ||   | : |:| ||:  |::|||||
Db       1 MEFWGVEVKNGKPLHLDPGLDRLVHISQVALGESKNNVTEPIQLYVTVGSDKLLIGTLSH 60

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGNAAK 120
           |  |||  ::| :: | ||||| |||:| |||     |:   ::: |      ||   |
Db      61 EKFPQLSTEIVLERNFALSHTWKNGSVFFSGYK-----PEDLIDDQLE-----AAGFKGK 110

Qy     121 -------AVAKPKAKPA-------EVKPAVDDEEDESDSDGMDEDDS------------ 153
                |  |  ||        :||   ||: | |: |  |:|||
Db     111 WGLLYPPAAPKSAAKQVNFQLPNEDVKAKQDDDADGSEEDSSDDDDSENSGDEEEEKVTA 170

Qy     154 -----------DGEDSEEEEPTPKKPASSKKRANE-TTPKAPVSAKKAKVAVTPQKTDEK 201
                     | ||   || |||||   |||: |  : | | | |||  |||||||
Db     171 ESDSEEDDSSDDEEDDSSEEETPKKPEEPKKRSAEPNSSKNPASNKKAKF-VTPQKTDSK 229

Qy     202 KK-----------------GGKAA-----NQSPKSASQVSCGSCKKTFNSGNALESHNK 238
           |                   ||         |:||||    | || :|| |   |:|| |
Db     230 KPHVHVATPHPSKQAGKNSGGGSTGETSKQQQTPKSAGAFGCKSCTRTFTSEMGLQSHTK 289

Qy     239 AKHAAA 244
           |||:||
Db     290 AKHSAA 295


RESULT   3
T04141
histone deacetylase (EC 3.5.1.-) HD2-p39, nucleolar - maize
C;Species: Zea mays (maize)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C;Accession: T04141
R;Lusser, A.; Brosch, G.; Loidl, A.; Haas, H.; Loidl, P.
Science 277, 88-91, 1997
A;Title: Identification of maize histone deacetylase HD2 as an acidic nucleolar phosphop
A;Reference number: Z15237; MUID:97349336
A;Accession: T04141
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-307 <LUS>
A;Cross-references: EMBL:U82815; NID:g2257755; PIDN:AAB63262.1; PID:g2257756
A;Experimental source: strain Cuzco 251
C;Keywords: hydrolase


  Query Match             37.7%;  Score 482;  DB 2;  Length 307;
  Best Local Similarity   39.0%;  Pred. No. 8.6e-24;
  Matches  121;  Conservative   38;  Mismatches   79;  Indels   72;  Gaps   11;

Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
           |||||:||| |   |   | | ::|:|||:|||  :|| :   ::||: :| | :||||
Db       1 MEFWGLEVKPGSTVKCEPGYGFVLHLSQAALGE--SKKSDNALMYVKIDDQKLAIGTLSV 58

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSE-------EEEEEEEEV 113
           :  | :  ||:|||||||||||    ||:| |||      | : |       |:|:||  |
Db      59 DKNPHIQFDLIFDKEFELSHTSKTTSVFFTGYKV----EQPFEEDEMDLDSEDEDEELNV 114

Qy     114 PA----GNA--------AKAVAKP-KAKPAEVKPAVDDEEDESDSDGMDEDDSD------ 154
           |    | |            |||| | |: |   |   ||: || ::|  |||:::|
Db     115 PVVKENGKADEKKQKSQEKAVAAPSKSSPDSKKSKDDDDSDEDETDDSDEDETDDSDEGL 174

Qy     155 ---------------GEDSEEEEPTPKKPASSKKR-ANETTPKAPVSAKKAKVAVTP--Q 196
                         :| ||: ||||||   ||| |  :  | |:| |||||| ||  |
Db     175 SSEEGDDDSSDEDDTSDDEEEDTPTPKKPEVGKKRPAESSVLKTPLSDKKAKVA-TPSSQ 233

Qy     197 KTDEKK---------------------KGGKAANQSPKSASQVSCGSCKKTFNSGNALES 235
           ||  ||                     |  |:   :|||   | |  | |:| |  ||::
Db     234 KTGGKKGAAVHVATPHPAKGKTIVNNDKSVKSPKSAPKSGGSVPCKPCSKSFISETALQA 293

Qy     236 HNKAKHAAAK 245
           |::||  |::
Db     294 HSRAKMGASE 303


RESULT   4
A55320
immunophilin FKBP46 - fall armyworm
C;Species: Spodoptera frugiperda (fall armyworm)
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 02-Sep-2000
C;Accession: A55320
R;Alnemri, E.S.; Fernandes-Alnemri, T.; Pomerenke, K.; Robertson, N.M.; Dudley, K.; D
J. Biol. Chem. 269, 30828-30834, 1994
A;Title: FKBP46, a novel Sf9 insect cell nuclear immunophilin that forms a protein-ki
A;Reference number: A55320; MUID:95074110
A;Accession: A55320
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-412 <ALN>
A;Cross-references: GB:U15038; NID:g595844; PID:g595845
C;Superfamily: yeast peptidylprolyl isomerase FPR3; BKBP-type peptidylprolyl isomeras
C;Keywords: nucleus; phosphoprotein
F;324-371/Domain: BKBP-type peptidylprolyl isomerase homology <PPI>


  Query Match             17.5%;  Score 224;  DB 2;  Length 412;
  Best Local Similarity   30.1%;  Pred. No. 2.9e-07;
  Matches   82;  Conservative   31;  Mismatches   91;  Indels   68;  Gaps   10;

Qy       3 FWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLSTEN 62
           |||: ::  |  |   |:    |:|||:: :      :   : | |  :| :: ||
Db       2 FWGLIMEPNKRYTQVVEKP--FHISQAAM-DISTGDNDPCQVMVVVDGKNFLVCTLQKGK 58

Qy      63 IPQLFCDLVFDKEFELSH-TWGKGSVYFVGYKTPNIEPQGYSEEEEEEEEE-------VP 114
           | |:  || |     :|  | || :|:  ||  |  |     ||| |||||       ||
Db      59 IIQVPLDLYFKSGDSVSFLTNGKCNVHLTGYLDPEFEEDLEDEEEAEEEEEEEEAPPLVP 118

Qy     115 AGNAAK------AVAKPKAKP-----------AEVKPAVDDEE---------------- 140
           | |  |       | |  ||||           ||     |||:
Db     119 AKNKRKLENANDATANKKAKPDKKAGKNSAPAAESDSDDDDEDQLQKFLDGEDIDTDEND 178

Qy     141 -----------DESDSDGMDED----DSDGEDSEEEEPTPKK----PASSKKRANETTPK 181
                      |:|| :  |||    : | || ||||  |||    ||: :    :|: :
Db     179 ESFKMNTSAEGDDSDEEDDDEDEEDEEDDDDEDDEEEEEAPKKKKKQPAAEQDSTLDTSKE 238

Qy     182 APVSAKKAKVAVTPQKTDEKKKGGKAANQSPK 213
           :   :| :|     ||   |||  : | | |:
Db     239 SVDMSKLSK----SQKRRLKKKLQQQAKQQPQ 266


RESULT   5
JC4090
FK506-binding 39k protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 23-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 21-Jul-2000
C;Accession: JC4090
R;Theopold, U.; Zotto, L.D.; Hultmark, D.
Gene 156, 247-251, 1995
```

```
A;Title: FKBP39, a Drosophila member of a family of proteins that bind the immunosuppres
A;Reference number: JC4090; MUID:95278752
A;Accession: JC4090
A;Molecule type: mRNA
A;Residues: 1-357 <THE>
A;Cross-references: EMBL:Z46894; NID:g600423; PIDN:CAA86996.1; PID:g600424
C;Genetics:
A;Gene: fkbp39
A;Cross-references: FlyBase:FBgn0013269
C;Superfamily: yeast peptidylprolyl isomerase FPR3; BKBP-type peptidylprolyl isomerase h
F;89-99/Region: acidic
F;252-357/Region: FK506 binding #status predicted
F;269-316/Domain: BKBP-type peptidylprolyl isomerase homology <PPI>


  Query Match             14.7%;  Score 188.5;  DB 2;  Length 357;
  Best Local Similarity   25.1%;  Pred. No. 4.6e-05;
  Matches   71;  Conservative   43;  Mismatches  110;  Indels   59;  Gaps   10;

Qy       3 FWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLSTEN 62
           |||: :|  :  : |  :    |:|  :|      ||:   |::    |  :: |: |:
Db       4 FWGLNMKPERKYSQTIIKS--FHISGVAL-----DKGQEAKLYLAAEKQEYIVATV-TKA 55

Qy      63 IPQLFCDLVFDK-EFELSHTWGKGSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGNAAKA 121
           |||:  || | | :  : :| | ||  :||           |::|:::::   |   :
Db      56 IPQVALDLNFSKGDRIMFYTAGDASVSLLGYL------HDIDSEDDEDDDQMTIENLLNS 109

Qy     122 VAKPKAKPA-----EVKPAVDDEED------------------ESDSDGMD-------- 149
            |   :| :     | :   :||||                    | | | :|
Db     110 KAIKNSKKSEDDEDENESGEEDEEDTDDDSQIIEEYESFLENGEEEDDDDVDEDNEESGE 169

Qy     150 EDDSDGEDSEEEEPTPKKP-------ASSKKRANETTPKAPVSAKKAKVAVTPQKTDEKK 202
           ||: | :||| ||  ||:         ||:||   |    |     | : |   ||  ||
Db     170 EDEQDSDDSEAEEEQPKRKVAKLSPGASAKKSGKEQNGVA-----KKEEAKQQQKKKEKP 224

Qy     203 KGGKAANQSPKSASQVSCGSCKKTFNSGNALESHNKAKHAAAK 245
           :  |   :: : | |       :|   |  :       |   ||
Db     225 EAKKEQPKAKEPAKQQPASKDPRTITGGVKIVDQVVGKGEEAK 267


RESULT   6
A29681
nucleophosmin - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 13-Aug-1999
C;Accession: A29681
R;Schmidt-Zachmann, M.S.; Huegle-Doerr, B.; Franke, W.W.
EMBO J. 6, 1881-1890, 1987
A;Title: A constitutive nucleolar protein identified as a member of the nucleoplasmin fa
A;Reference number: A29681; MUID:88004394
A;Accession: A29681
A;Molecule type: mRNA
A;Residues: 1-299 <SCH>
A;Cross-references: GB:X05496; NID:g64924; PIDN:CAA29046.1; PID:g64925
C;Superfamily: nucleophosmin
C;Keywords: molecular chaperone; nucleus


  Query Match             13.3%;  Score 170.5;  DB 2;  Length 299;
  Best Local Similarity   31.5%;  Pred. No. 0.00053;
  Matches   56;  Conservative   20;  Mismatches   67;  Indels   35;  Gaps    7;

Qy      83 GKGSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGNAAKAVAKPKAKPAEVKPAVDDEEDE 142
           | | ||  |     :|    |::|:|| |  |  || :      :|    |  :::||::
Db     107 GSGPVYVSGQHLVALEDLESSDDEDEEHEPSPK-NAKRIAPDSASKVPRKKTRLEEEEED 165

Qy     143 SD-SDGMDEDDSDGEDSEEEEPTP-KKPASSKKRA---------------NETTPKAPVS 185
           ||  |  |||| | :| |||| || ||  |:| :|                 : ||| |
Db     166 SDEDDDDDEDDDDEDDDEEEEETPVKKTDSTKSKAAQKLNHNGKASALSTTQKTPKTPEQ 225

Qy     186 AKK-------AKVAVTPQKTDEKK--------KGGKAANQSPKSASQVSCGSCKKTFN 228
           |         |   ||  ::| |         ||        | |: |   :| :| | |
Db     226 KGKQDTKPQTPKTPKTPLSSEEIKAKMQTYLEKGNVLPKVEVKFANYVK--NCFRTEN 281


RESULT   7
DNCHNL
nucleolin - chicken
N;Alternate names: nucleolar protein C23
C;Species: Gallus gallus (chicken)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 22-Jun-1999
C;Accession: S08414; S10766; A32725; I50397; B30099
R;Maridor, G.; Nigg, E.A.
Nucleic Acids Res. 18, 1286, 1990
A;Title: cDNA sequences of chicken nucleolin/C23 and NO38/B23, two major nucleolar pr
A;Reference number: S08414; MUID:90206792
A;Accession: S08414
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-694 <MAR>
A;Cross-references: EMBL:X17199; NID:g63710; PIDN:CAA35060.1; PID:g63711
R;Maridor, G.; Krek, W.; Nigg, E.A.
Biochim. Biophys. Acta 1049, 126-133, 1990
A;Title: Structure and developmental expression of chicken nucleolin and NO38: coordi
A;Reference number: S10766; MUID:90304215
A;Accession: S10766
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-694 <MAR2>
R;Peter, M.; Nakagawa, J.; Doree, M.; Labbe, J.C.; Nigg, E.A.
Cell 60, 791-801, 1990
A;Title: Identification of major nucleolar proteins as candidate mitotic substrates o
A;Reference number: A32725; MUID:90182668
A;Accession: A32725
A;Molecule type: protein
A;Residues: 56-62;63-109;197-214 <PET>
R;Borer, R.A.; Lehner, C.F.; Eppenberger, H.M.; Nigg, E.A.
Cell 56, 379-390, 1989
A;Title: Major Nucleolar Proteins Shuttle between Nucleus and Cytoplasm.
A;Reference number: I50397; MUID:89119560
A;Accession: I50397
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 407-418,'R',420-519,'T',521-694 <BOR>
A;Cross-references: GB:M21791; NID:g212411; PIDN:AAA48983.1; PID:g212412
C;Comment: Phosphorylation of this protein by cdc2 kinase may contribute to the reorg
C;Superfamily: nucleolin; ribonucleoprotein repeat homology
C;Keywords: DNA binding; duplication; nucleolus; nucleus; phosphoprotein; RNA binding
F;1-247/Domain: acidic <ACI>
F;54-91/Region: 7-residue repeats (T-P-A-K-K-A-[A/V])
F;254-262/Region: nuclear location signal
F;267-275/Region: nuclear location signal
F;282-347/Domain: ribonucleoprotein repeat homology <RRM1>
F;283-361/Domain: RNA binding #status predicted <RNA1>
F;372-435/Domain: ribonucleoprotein repeat homology <RRM2>
F;373-448/Domain: RNA binding #status predicted <RNA2>
F;462-525/Domain: ribonucleoprotein repeat homology <RRM3>
F;463-538/Domain: RNA binding #status predicted <RNA3>
F;554-618/Domain: ribonucleoprotein repeat homology <RRM4>
F;555-631/Domain: RNA binding #status predicted <RNA4>
F;632-694/Domain: glycine/arginine-rich <GRR>
F;56,63,70,77,85/Binding site: phosphate (Thr) (covalent) (by cdc2 kinase) #status pr


  Query Match             13.3%;  Score 170;  DB 1;  Length 694;
  Best Local Similarity   33.9%;  Pred. No. 0.0013;
  Matches   60;  Conservative   16;  Mismatches   61;  Indels   40;  Gaps    8;

Qy      84 KGSVYFVGYKTPNIEPQGYSEEEEEEE-----------EEVPAGNAAKAVAKP---KAK 128
           | |   |  | | : |     ||||||             |:          | ||   ||
Db     151 KKSAAAVPAKKPAVVPAKQESEEEEEEDDEEEDEEDDESEDEAMDTTPAPVKKPTPAKAT 210

Qy     129 PAEVKPAVDDEEDESDSDGMDEDDSDGEDSEEEEPTPKKPAS------SKKRANETTPKA 182
           ||: |   :|||||  | |  ||:| | |: :|||   :||          |: ||:: |:|
Db     211 PAKAKAESEDEEDEEDED-EDEEDEDDEEEDEEESEDEKPVKEAPGKRKKEMANKSAPEA 269
```

```
Qy    183 -------PVSAKKAKVA-VTPQKTDE------KKKGGKAANQSPKSASQVSCGSCKK 225
                  | ||    |  :|| |  |      |:  ||   |     |:|  || |:
Db    270 KKKKTETPASAFSLFVKNLTPTKDYEELRTAIKEFFGKKNLQ----VSEVRIGSSKR 322


RESULT   8
DNCHFM
nucleophosmin - chicken
N;Alternate names: nucleolar protein B23; nucleolar protein NO38; numatrin
C;Species: Gallus gallus (chicken)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 22-Jun-1999
C;Accession: S08415; S10767; I50398; I50399; A30099
R;Maridor, G.; Nigg, E.A.
Nucleic Acids Res. 18, 1286, 1990
A;Title: cDNA sequences of chicken nucleolin/C23 and NO38/B23, two major nucleolar prote
A;Reference number: S08414; MUID:90206792
A;Accession: S08415
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-294 <MAR>
A;Cross-references: EMBL:X17200; NID:g63704; PIDN:CAA35061.1; PID:g63705
R;Maridor, G.; Krek, W.; Nigg, E.A.
Biochim. Biophys. Acta 1049, 126-133, 1990
A;Title: Structure and developmental expression of chicken nucleolin and NO38: coordinat
A;Reference number: S10766; MUID:90304215
A;Accession: S10767
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-294 <MAR2>
R;Borer, R.A.; Lehner, C.F.; Eppenberger, H.M.; Nigg, E.A.
Cell 56, 379-390, 1989
A;Title: Major Nucleolar Proteins Shuttle between Nucleus and Cytoplasm.
A;Reference number: I50397; MUID:89119560
A;Accession: I50398
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-97 <BOR>
A;Cross-references: GB:J03164; NID:g212453; PIDN:AAA48989.1; PID:g212456
A;Accession: I50399
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 261-294 <BO2>
A;Cross-references: GB:J03165; NID:g212454; PIDN:AAA48990.1; PID:g212457
C;Superfamily: nucleophosmin
C;Keywords: nucleus; phosphoprotein
F;122-133/Region: glutamic acid-rich
F;153-159/Region: nuclear location signal
F;161-186/Region: aspartic acid/glutamic acid-rich
F;190-197/Region: nuclear location signal
F;127/Binding site: phosphate (Ser) (covalent) #status predicted


  Query Match             13.2%;  Score 169;  DB 1;  Length 294;
  Best Local Similarity   34.9%;  Pred. No. 0.00065;
  Matches   51;  Conservative   24;  Mismatches   57;  Indels   14;  Gaps    7;

Qy     76 FELSHTWGKGSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGNAA--KAVAKPKAKPAEVK 133
          | |    | | ||  |     :| :    | |:|||:   |||:  :  :   ||  : |
Db    100 FVLRLKCGSGPVYVSGQHLVALEEE---PESEDEEEDTKIGNASTKRPASGGGAKTPQKK 156

Qy    134 PAV--DDEEDESDSDGMDEDDSDGEDSEEEEPTP-KKPASSKKRANETTPKAPVSAKKAK 190
          | :  |||:|: | |  ::|: | :| |||  || ||||  :: | :   ||  : | :|
Db    157 PKLSEDDEDDDEDEDDDEDDEDLDDDEEEIKTPMKKPA--REPAGKNMQKAKQNGKDSK 214

Qy    191 VAVTP---QKTDEKKKGGKAANQSPK 213
            : ||    || : ||      ::||
Db    215 PS-TPASKTKTPDSKKDKSLTPKTPK 239


RESULT   9
A27441
nucleolin - Chinese hamster
N;Alternate names: nonchromosomal protein C23; nucleolar protein C23
C;Species: Cricetulus griseus (Chinese hamster)
C;Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 23-Jul-1999
C;Accession: A27441; A24808; A24070
R;Lapeyre, B.; Bourbon, H.; Amalric, F.
Proc. Natl. Acad. Sci. U.S.A. 84, 1472-1476, 1987
A;Title: Nucleolin, the major nucleolar protein of growing eukaryotic cells: an unusu
A;Reference number: A27441; MUID:87175501
A;Accession: A27441
A;Molecule type: mRNA
A;Residues: 1-713 <LAP>
R;Lapeyre, B.; Amalric, F.; Ghaffari, S.H.; Venkatarama Rao, S.V.; Dumbar, T.S.; Olso
J. Biol. Chem. 261, 9167-9173, 1986
A;Title: Protein and cDNA sequence of a glycine-rich, dimethylarginine-containing reg
A;Reference number: A24808; MUID:86250857
A;Accession: A24808
A;Molecule type: mRNA
A;Residues: 522-541,'G',543-544,'R',546-713 <LA2>
A;Cross-references: GB:M13757; NID:g191004; PIDN:AAA36961.1; PID:g387047
R;Lapeyre, B.; Caizergues-Ferrer, M.; Bouche, G.; Amalric, F.
Nucleic Acids Res. 13, 5805-5816, 1985
A;Title: Cloning of cDNA encoding a 100 kDa nucleolar protein (nucleoline) of Chinese
A;Reference number: A24070; MUID:85297777
A;Accession: A24070
A;Molecule type: mRNA
A;Residues: 179-238 <LA3>
C;Superfamily: nucleolin; ribonucleoprotein repeat homology
C;Keywords: DNA binding; nucleus; phosphoprotein
F;308-373/Domain: ribonucleoprotein repeat homology <RRM1>
F;394-456/Domain: ribonucleoprotein repeat homology <RRM2>
F;486-549/Domain: ribonucleoprotein repeat homology <RRM3>
F;572-636/Domain: ribonucleoprotein repeat homology <RRM4>


  Query Match             13.0%;  Score 166.5;  DB 2;  Length 713;
  Best Local Similarity   33.8%;  Pred. No. 0.0023;
  Matches   48;  Conservative   20;  Mismatches   55;  Indels   19;  Gaps    4;

Qy     73 DKEFELSHTWGK-GSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGNAAKAV----AKPKA 127
          : |||     || | |       : : : :   |||||:|||       :|:    || |
Db    165 EDEFEPPVVKGKQGKVAAAAPASEDEDEEEDEEEEEEDEEEEDDSEEEEAMEITPAKGKK 224

Qy    128 KPAEVKPAV----------DDEEDESDSDGMDEDDSDGEDSEEEEPTPKKPASSKKRANE 177
           ||:| |             |:|||| : :  :|:: : |: ||||  | |||  |::
Db    225 APAKVVPVKAKNVAEEDDDEEEDEDEEDEEEEEDEEEEEEEEEEPVKPAPGKRKKEM 284

Qy    178 TTPKAPVSAKKAKV----AVTP 195
          |  |    ||| ||    : ||
Db    285 TKQKEVPEAKKQKVEGSESTTP 306


RESULT  10
I51618
nucleolar phosphoprotein - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 24-Sep-1999
C;Accession: I51618; S57757
R;Cairns, C.; McStay, B.
J. Cell Sci. 108, 3339-3347, 1995
A;Title: Identification and cDNA cloning of a Xenopus nucleolar phosphoprotein, xNopp
A;Reference number: I51618; MUID:96019267
A;Accession: I51618
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-990 <CAI>
A;Cross-references: EMBL:X88927; NID:g895920; PIDN:CAA61368.1; PID:g895921
C;Genetics:
A;Gene: xNopp180
C;Superfamily: nucleolus-cytoplasm shuttle phosphoprotein
C;Keywords: phosphoprotein


  Query Match             12.9%;  Score 165.5;  DB 2;  Length 990;
  Best Local Similarity   30.7%;  Pred. No. 0.0037;
```

```
  Matches   55;  Conservative   17;  Mismatches   64;  Indels   43;  Gaps    7;

Qy     103 SEEEEEEEEEVPAGN--AAKAVAKPKAKPAEVKPAVDDEEDESDSDGMDEDDSD------ 154
           |  ||| : : | |   |||| | ||  |  |          ||| | ||
Db     703 SSSEEEGKSKQPTGKSPAAKATAPPKKNPVAVNKDKPSSSSSSDSSGDDEKQKPKQAAAA 762

Qy     155 ---GEDSEEEEPTPKKPASSK------------KRANE-------TTPKAPVSAKKAKV 191
               : ::  :||||| |||              |: |        |||||  :|||
Db     763 KDVKQGAKAAKPTPKKAASSSSEDSSSDEDVSKAKKTNTAVSKSPVTTPKAVPAAKKESS 822

Qy     192 AVTPQKTDEKKKG-----GKAANQSPKSASQVSCGSCKKTFNSGNALESHNKAKHAAAK 245
           : :    |||: |      | || :||:|:      | :  :|    |   ||   : |
Db     823 SESSDSEDEKQGGKNTSTTKIANSTPKAAA----AECSEESSSS---EDEGKANGTSGK 874


RESULT  11
T10215
hypothetical protein T30C3.20 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 15-Oct-1999
C;Accession: T10215
R;Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z16897
A;Accession: T10215
A;Molecule type: DNA
A;Residues: 1-487 <BEV>
A;Cross-references: EMBL:AL079350; GSPDB:GN00062; ATSP:T30C3.20
A;Experimental source: cultivar Columbia; BAC clone T30C3
C;Genetics:
A;Gene: ATSP:T30C3.20
A;Map position: 4
A;Introns: 5/1; 39/3; 122/2; 170/1; 271/2; 327/2; 401/3; 431/1; 445/1; 461/2
C;Superfamily: BKBP-type peptidylprolyl isomerase homology
F;399-446/Domain: BKBP-type peptidylprolyl isomerase homology <PPI>


  Query Match             12.8%;  Score 164;  DB 2;  Length 487;
  Best Local Similarity   26.5%;  Pred. No. 0.0022;
  Matches   83;  Conservative   33;  Mismatches   87;  Indels  110;  Gaps   16;

Qy       1 MEFW----------GIEVKSGKPVTVTP--EEGILIHVSQASLGECKNKKGEFVPLHVKV 48
           | ||          |:||| |||    |  |:|  |||:||:||   :|:   :     :
Db       1 MGFWGSDFVLLTYAGLEVKPGKPQAYNPKNEQG-KIHVTQATLGTGLSKEKSVI--QCSI 57

Qy      49 GNQ-NLVLGTLSTENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSEEEE 107
           |::  : | :|    |     :| || : |         | |      :|   |: |  :
Db      58 GDKAPIALCSLLPNKIECCPLNLEFDDDDE--------PVEFTVTGDRSIHLSGFLEYYQ 109

Qy     108 EEEEEVPAGNAAKAVAKPKAKPAEVKPAVDDEEDESDSDGMD-----EDDSDGEDSEEEE 162
           ::|:                          | | || ||||:|     ||||   ||||:|
Db     110 DDED------------------------DYEHDEDDSDGIDVGESEEDDSCEYDSEEDE 144

Qy     163 -------------------PTPK------------KPASSKKRANETTPKAPVS----AK 187
                               ||             |||   | | :|  |:  |    ||
Db     145 QLDEFEDFLDSNLERYRNAAAPKSGVIIEEIEDEEKPAKDNK-AKQTKKKSQASEGENAK 203

Qy     188 KAKVAV------TPQKTDEKKKG-----GKAANQSPKSASQV-----SCGSCK----KTF 227
           |  ||:        :  || : |     ||::     |  ::       || |    |
Db     204 KQIVAIEGAHVPVLESEDEDEDGLPIPKGKSSEVENASGEKMVVDNDEQGSNKKRKAKAA 263

Qy     228 NSGNALESHNKAK 240
             :  || ||:|
Db     264 EQDDGQESANKSK 276


RESULT  12
I51116
NF-180 - sea lamprey
C;Species: Petromyzon marinus (sea lamprey)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C;Accession: I51116
R;Jacobs, A.J.; Kamholz, J.; Selzer, M.E.
Brain Res. Mol. Brain Res. 29, 43-52, 1995
A;Title: The single lamprey neurofilament subunit (NF-180) lacks multiphosphorylation
A;Reference number: I51116; MUID:95287814
A;Accession: I51116
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1110 <JAC>
A;Cross-references: EMBL:U19361; NID:g632548; PIDN:AAA80106.1; PID:g632549
C;Superfamily: neurofilament triplet H protein


  Query Match             12.3%;  Score 157;  DB 2;  Length 1110;
  Best Local Similarity   32.5%;  Pred. No. 0.014;
  Matches   41;  Conservative   19;  Mismatches   42;  Indels   24;  Gaps    3;

Qy     105 EEEEEEEEVP---------AGNAAKAVAKPK------------AKPAEVK--PAVDDEE 140
           |:||| || |         |   ||| ||||            ||||: |  || ::||
Db     854 EKEEEPEESPTEEEPKKPAAAKPAKAPAKPKPAPKAEAEEKPEPAKPAQAKPAPAAEEEE 913

Qy     141 DESDSDGMDEDDSDGEDSEEEEPTPKKPASSKKRANETTPKAPVSAKKAKVAVTPQKTDE 200
           || : |  :|:: :    |: :|   ||| :|:  :|  |       | | |   :: ::
Db     914 DEKEDDEEEEEEVEEVKPEDAKPVKSKPAPAKEEEDEPKPAKQPPKPKRKPARPKEEPED 973

Qy     201 KKKGGK 206
           | :  |
Db     974 KAEPAK 979


RESULT  13
A32915
nucleophosmin - human
N;Alternate names: nucleolar phosphoprotein B23; numatrin
C;Species: Homo sapiens (man)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C;Accession: A33423; A32915; A30137; S06926; B26080; A26080
R;Zhang, X.T.; Thomis, D.C.; Samuel, C.E.
Biochem. Biophys. Res. Commun. 164, 176-184, 1989
A;Title: Isolation and characterization of a molecular cDNA clone of a human mRNA fro
A;Reference number: A33423; MUID:90026373
A;Accession: A33423
A;Molecule type: mRNA
A;Residues: 1-294 <ZHA>
A;Cross-references: GB:M26697; NID:g189311; PIDN:AAA36385.1; PID:g189312
R;Li, X.; McNeilage, L.J.; Whittingham, S.
Biochem. Biophys. Res. Commun. 163, 72-78, 1989
A;Title: The nucleotide sequence of a human cDNA encoding the highly conserved nucleo
A;Reference number: A32915; MUID:89374300
A;Accession: A32915
A;Molecule type: mRNA
A;Residues: 1-294 <LIX>
A;Cross-references: GB:M28699; NID:g557545; PIDN:AAA58386.1; PID:g557546
R;Chan, W.Y.; Liu, Q.R.; Borjigin, J.; Busch, H.; Rennert, O.M.; Tease, L.A.; Chan, P
Biochemistry 28, 1033-1039, 1989
A;Title: Characterization of the cDNA encoding human nucleophosmin and studies of its
A;Reference number: A30137; MUID:89229105
A;Accession: A30137
A;Molecule type: mRNA
A;Residues: 1-294 <CHAW>
A;Cross-references: GB:M23613; NID:g189271; PIDN:AAA36380.1; PID:g189272
R;Hale, T.K.; Mansfield, B.C.
Nucleic Acids Res. 17, 10112, 1989
A;Title: Nucleotide sequence of a cDNA clone representing a third allele of human pro
A;Reference number: S06926; MUID:90098787
A;Accession: S06926
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 15-294 <HAL>
A;Cross-references: EMBL:X16934; NID:g32029; PIDN:CAA34809.1; PID:g825671
A;Experimental source: clone hL6T
R;Chan, P.K.; Chan, W.Y.; Yung, B.Y.M.; Cook, R.G.; Aldrich, M.B.; Ku, D.; Goldknopf,
J. Biol. Chem. 261, 14335-14341, 1986
A;Title: Amino acid sequence of a specific antigenic peptide of protein B23.
A;Reference number: A92562; MUID:87033628
```

```
A;Accession: B26080
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 'P',214-226 <CHA1>
A;Cross-references: GB:M31004; GB:M12775; NID:g190239
A;Experimental source: clone hpB2
A;Accession: A26080
A;Molecule type: mRNA
A;Residues: 'P',214-215,'S',217-218,'SSS',222-294 <CHA2>
A;Cross-references: GB:J02590; NID:g190237; PIDN:AAA36473.1; PID:g190238
C;Genetics:
A;Gene: GDB:NPM1
A;Cross-references: GDB:138455; OMIM:164040
A;Map position: 5q35-5q35
C;Superfamily: nucleophosmin
C;Keywords: nucleolus; nucleus; phosphoprotein
F;120-132/Region: glutamic acid-rich
F;152-158/Region: nuclear location signal
F;161-188/Region: aspartic acid/glutamic acid-rich
F;191-198/Region: nuclear location signal
F;125/Binding site: phosphate (Ser) (covalent) #status predicted


  Query Match             12.1%;  Score 155;  DB 1;  Length 294;
  Best Local Similarity   28.3%;  Pred. No. 0.0051;
  Matches   43;  Conservative   25;  Mismatches   56;  Indels   28;  Gaps    5;

Qy      83 GKGSVYFVGYKTPNIEPQGYSEEEEEEE---------EEVPAGNAAKAVAKPKAKPAEVK 133
           | | |:  |     :|    ||:||||:          | |       :|   |  ::
Db     105 GSGPVHISGQHLVAVEEDAESEDEEEEDVKLLSISGKRSAPGGG-----SKVPQKKVKLA 159

Qy     134 PAVDDEEDESDSDGMDEDDSDGEDSEEEEPTPKK------PASSKKRANETTPKAPVSAK 187
             ||::|: : |  |:|| | :| | ||  | |      || : :::|:       : |
Db     160 ADEDDDDDDEEDDDEDDDDDDFDDEEAEEKAPVKKSIRDTPAKNAQKSNQ-------NGK 212

Qy     188 KAKVAVTPQ-KTDEKKKGGKAANQSPKSASQV 218
           :| : ||: |  |  |  :   ::||  | |
Db     213 DSKPSSTPRSKGQESFKKQEKTPKTPKGPSSV 244


RESULT  14
T09648
nucleolin homolog nuM1 - alfalfa
N;Alternate names: nuM1 protein
C;Species: Medicago sativa (alfalfa)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C;Accession: T09648
R;Boegre, L.; Jonak, C.; Mink, M.; Meskiene, I.; Traas, J.; Ha, D.T.C.; Swoboda, I.; Pla
Plant Cell 8, 417-428, 1996
A;Title: Developmental and cell cycle regulation of Alfalfa nucMs1 a plant homolog of th
A;Reference number: Z16796; MUID:96361876
A;Accession: T09648
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-635 <BOE>
A;Cross-references: EMBL:X88845; NID:g1279562; PIDN:CAA61298.1; PID:g1279563
C;Genetics:
A;Gene: nuM1
C;Superfamily: nucleolin; ribonucleoprotein repeat homology


  Query Match             12.0%;  Score 153.5;  DB 2;  Length 635;
  Best Local Similarity   28.5%;  Pred. No. 0.014;
  Matches   47;  Conservative   17;  Mismatches   60;  Indels   41;  Gaps    5;

Qy     105 EEEEEEEEVPAGNAAKAVAKPKAKPAEVKPAVDDEEDESDSDG--------------MDE 150
           |  :|||:: ||  |:| |: |  | |        |||  : | |               |
Db     221 ESSDEEDKKPAAKASKNVSAPTKKAASSSDEESDEESDEDEDAKPVSKPAAVAKKSKKDS 280

Qy     151 DDSDGEDSEEEEPTPKKPASSKKRANETTPKAPVSAKKAKV--------------AVTPQ 196
             ||| || :      ||| :|||  :|:  :    |  :                |
Db     281 SDSDDEDDDSSSDEDKKPVASKKEVSESESDSSDDDHKMNIDKDSSDSDESEEESEDEPL 340

Qy     197 KTDEKK----------KGGKAANQSPKSASQVSCGSCKKTFNSGN 231
           || :||           | || |  :| : :: | ||  ||   ||
Db     341 KTPQKKIKDVEMVDAGKSGKKAPNTPATPNETS-GS--KTLFVGN 382


RESULT  15
DNMS
nucleolin - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 22-Jun-1999
C;Accession: A29958; A40769; A56240; I84688
R;Bourbon, H.M.; Lapeyre, B.; Amalric, F.
J. Mol. Biol. 200, 627-638, 1988
A;Title: Structure of the mouse nucleolin gene. The complete sequence reveals that ea
A;Reference number: A29958; MUID:88316930
A;Accession: A29958
A;Molecule type: DNA
A;Residues: 1-707 <BOU>
A;Cross-references: GB:X07699; NID:g53453; PIDN:CAA30538.1; PID:g53454
R;Pasternack, M.S.; Bleier, K.J.; McInerney, T.N.
J. Biol. Chem. 266, 14703-14708, 1991
A;Title: Granzyme A binding to target cell proteins. Granzyme A binds to and cleaves
A;Reference number: A40769; MUID:91317840
A;Accession: A40769
A;Molecule type: protein
A;Residues: 2-20,'X',22-24 <PAS>
R;Yang, T.H.; Tsai, W.H.; Lee, Y.M.; Lei, H.Y.; Lai, M.Y.; Chen, D.S.; Yeh, N.H.; Lee
Mol. Cell. Biol. 14, 6068-6074, 1994
A;Title: Purification and characterization of nucleolin and its identification as a t
A;Reference number: A56240; MUID:94344117
A;Accession: A56240
A;Molecule type: protein
A;Residues: 2-19;558-567 <YAN>
R;Bourbon, H.
Gene 68, 73-84, 1988
A;Title: Sequence and structure of the nucleolin promoter in rodents: Characterizatio
A;Reference number: I48118; MUID:89121496
A;Accession: I84688
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-44 <RES>
A;Cross-references: GB:M22089; NID:g200111; PIDN:AAA39841.1; PID:g554246
C;Comment: This housekeeping protein is involved in the synthesis, packaging, and mat
C;Genetics:
A;Introns: 6/3; 45/3; 211/1; 273/1; 302/1; 349/2; 391/1; 432/3; 484/1; 525/2; 566/1;
C;Superfamily: nucleolin; ribonucleoprotein repeat homology
C;Keywords: DNA binding; duplication; nucleus; phosphoprotein; RNA binding; transcrip
F;310-375/Domain: ribonucleoprotein repeat homology <RRM1>
F;311-316/Region: RNA-binding RNP2 motif
F;349-356/Region: RNA-binding RNP1 motif
F;396-458/Domain: ribonucleoprotein repeat homology <RRM2>
F;397-402/Region: RNA-binding RNP2 motif
F;431-438/Region: RNA-binding RNP1 motif
F;488-551/Domain: ribonucleoprotein repeat homology <RRM3>
F;489-494/Region: RNA-binding RNP2 motif
F;524-531/Region: RNA-binding RNP1 motif
F;570-634/Domain: ribonucleoprotein repeat homology <RRM4>
F;571-576/Region: RNA-binding RNP2 motif
F;607-614/Region: RNA-binding RNP1 motif


  Query Match             12.0%;  Score 153.5;  DB 1;  Length 707;
  Best Local Similarity   24.1%;  Pred. No. 0.015;
  Matches   61;  Conservative   31;  Mismatches   94;  Indels   67;  Gaps    6;

Qy       6 IEVKSGKPVTVTPEEGILI----------------------------HVSQASLGECK 35
           |  | ||  | || : :::                            ::: | :
Db      48 IPQKKGKKATTTPAKKVVVSQTKKAAVPTPAKKAAVTPGKKAVATPAKKNITPAKVIPTP 107

Qy      36 NKKG-----EFVPLHVKVGNQNLVLGTLSTENIPQLFCDLVFDKEFELSHTWGKGSVYFV 90
             |||      ||   | |      |  : :|  :   |   |:| |      :
Db     108 GKKGAAQAKALVPTPGKKGAATPAKGAKNGKNAKKEDSDEDEDEEDEDDSDEDEDDEEED 167
```

```
Qy      91 GYKTPNI----------------EPQGYSEEEEEEEEEVPAGNAAKAV-----AKPKAK 128
           :: | :                 |     |::||:::|    :: : |     || |
Db     168 EFEPPIVKGVKPAKAAPAAPASEDEEDDEDEDDEEDDDEEEEDDSEEEVMEITTAKGKKT 227

Qy     129 PAEVKPAV---------DDEEDESDSDGMD-EDDSDGEDSEEEEPTPKKPASSKKRANET 178
           ||:| |           |:|||| | |  | |:| : :| ||||  | | |  |::   |
Db     228 PAKVVPMKAKSVAEEEDDEEEDEDDEDEDDEEEDDEDDDEEEEEEEPVKAAPGKRKKEMT 287

Qy     179 TPKAPVSAKKAKV 191
           |    ||| ||
Db     288 KQKEAPEAKKQKV 300


Search completed: May  2, 2002, 22:07:30
Job time: 6678 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: May 2, 2002, 22:07:34 ; Search time 36.81 Seconds
(without alignments)
257.710 Million cell updates/sec

Title: US-09-645-337-6
Perfect score: 1279
Sequence: 1 MEFWGIEVKSGKPVTVTPEE.....TFNSGNALESHNKAKHAAAK 245

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	224	17.5	412	1	FKB4_SPOFR	Q26486 spodoptera
2	191.5	15.0	357	1	FKB4_DROME	P54397 drosophila
3	170.5	13.3	299	1	NPM_XENLA	P07222 xenopus lae
4	170	13.3	694	1	NUCL_CHICK	P15771 gallus gall
5	169	13.2	294	1	NPM_CHICK	P16039 gallus gall
6	166.5	13.0	713	1	NUCL_MESAU	P08199 mesocricetu
7	155	12.1	294	1	NPM_HUMAN	P06748 homo sapien
8	153.5	12.0	411	1	MP62_LYTPI	P91753 lytechinus
9	153.5	12.0	706	1	NUCL_MOUSE	P09405 mus musculu
10	153	12.0	706	1	NUCL_HUMAN	P19338 homo sapien
11	151.5	11.8	712	1	NUCL_RAT	P13383 rattus norv
12	150.5	11.8	281	1	HS32_DICDI	P54658 dictyosteli
13	149.5	11.7	644	1	NFM_RABIT	P54938 oryctolagus
14	149	11.6	292	1	NPM_RAT	P13084 rattus norv
15	145.5	11.4	704	1	NP14_RAT	P41777 rattus norv
16	144	11.3	292	1	NPM_MOUSE	Q61937 mus musculu
17	139.5	10.9	411	1	FKB3_YEAST	P38911 saccharomyc
18	138	10.8	699	1	NP14_HUMAN	Q14978 homo sapien
19	137	10.7	590	1	YNW7_YEAST	P53863 saccharomyc
20	136	10.6	650	1	NUCL_XENLA	P20397 xenopus lae
21	131.5	10.3	731	1	MR11_COPCI	Q9uvn9 coprinus ci
22	130	10.2	1982	1	CHDM_DROME	O97159 drosophila
23	129.5	10.1	1085	1	IFH1_YEAST	P39520 saccharomyc
24	128.5	10.0	1411	1	TCOF_HUMAN	Q13428 homo sapien
25	128	10.0	639	1	SSTP_CATRO	Q39601 catharanthu
26	128	10.0	891	1	MAZ3_SCHCO	P37937 schizophyll
27	127.5	10.0	361	1	FKB4_SCHPO	O74191 schizosacch
28	126.5	9.9	723	1	SSRP_DROME	Q05344 drosophila
29	126	9.9	217	1	KS1_HYDAT	P38978 hydra atten
30	126	9.9	226	1	BASP_BOVIN	P80724 bos taurus
31	126	9.9	450	1	CY1_PARDE	P13627 paracoccus
32	126	9.9	918	1	IF38_SCHPO	O14164 schizosacch
33	125.5	9.8	200	1	NUPL_XENLA	P05221 xenopus lae
34	125	9.8	618	1	ORC2_DROME	Q24168 drosophila
35	124.5	9.7	810	1	NFM_BOVIN	O77788 bos taurus
36	123.5	9.7	845	1	NFM_RAT	P12839 rattus norv
37	123.5	9.7	1359	1	ATRX_CAEEL	Q9u7e0 caenorhabdi
38	123	9.6	346	1	RS6_AEDAE	Q9u761 aedes aegyp
39	123	9.6	857	1	NFM_CHICK	P16053 gallus gall
40	122.5	9.6	677	1	UBF1_XENLA	P25979 xenopus lae
41	122.5	9.6	915	1	NFM_HUMAN	P07197 homo sapien
42	122	9.5	517	1	T2FA_HUMAN	P35269 homo sapien
43	122	9.5	1095	1	NEB1_RAT	O35867 rattus norv
44	121.5	9.5	1020	1	NFH_HUMAN	P12036 homo sapien
45	121.5	9.5	1087	1	NFH_MOUSE	P19246 mus musculu

ALIGNMENTS

```
RESULT   1
FKB4_SPOFR
ID   FKB4_SPOFR      STANDARD;      PRT;   412 AA.
AC   Q26486;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   46 kDa FK506-binding nuclear protein (Peptidyl-prolyl cis-trans
DE   isomerase) (PPiase) (EC 5.2.1.8).
OS   Spodoptera frugiperda (Fall armyworm).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Noctuoidea; Noctuidae; Amphipyrinae; Spodoptera.
OX   NCBI_TaxID=7108;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95074110; PubMed=7527037;
RA   Alnemri E.S., Fernandes-Alnemri T., Pomerenke K., Robertson N.M.,
RA   Dudley K., Dubois G.C., Litwack G.;
RT   "FKBP46, a novel Sf9 insect cell nuclear immunophilin that forms a
RT   protein-kinase complex.";
RL   J. Biol. Chem. 269:30828-30834(1994).
CC   -!- FUNCTION: BINDS TO, AND IS INHIBITED BY FK506 AND RAPAMYCIN. BINDS
CC       DOUBLE-STRANDED DNA IN VITRO. PPIASES ACCELERATE PROTEIN FOLDING.
CC   -!- CATALYTIC ACTIVITY: CIS-TRANS ISOMERIZATION OF PROLINE IMIDIC
CC       PEPTIDE BONDS IN OLIGOPEPTIDES.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- PTM: PHOSPHORYLATED BY A NUCLEAR KINASE IN THE PRESENCE OF MG2+
CC       AND ATP.
CC   -!- SIMILARITY: BELONGS TO THE FKBP-TYPE PPIASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U15038; AAA58962.1; -.
DR   HSSP; P20071; 1FKJ.
DR   InterPro; IPR001179; FKBP_PPIase.
DR   Pfam; PF00254; FKBP; 1.
DR   PROSITE; PS00453; FKBP_PPIASE_1; FALSE_NEG.
DR   PROSITE; PS00454; FKBP_PPIASE_2; 1.
DR   PROSITE; PS50059; FKBP_PPIASE_3; 1.
KW   Isomerase; Rotamase; Nuclear protein; DNA-binding; Phosphorylation.
FT   DOMAIN        91    112       ASP/GLU-RICH (HIGHLY ACIDIC).
FT   DOMAIN       120    145       LYS-RICH (BASIC).
FT   DOMAIN       152    216       ASP/GLU-RICH (HIGHLY ACIDIC).
FT   DOMAIN       219    302       LYS-RICH (BASIC).
FT   DOMAIN       324    412       PPIASE, FKBP-TYPE.
SQ   SEQUENCE   412 AA;  45810 MW;  F2A69159AEF4FE22 CRC64;


  Query Match             17.5%;  Score 224;  DB 1;  Length 412;
```

```
  Best Local Similarity   30.1%;  Pred. No. 1.5e-07;
  Matches   82;  Conservative   31;  Mismatches   91;  Indels   68;  Gaps   10;

Qy       3 FWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLSTEN 62
           |||: ::  |  |   |:    |:|||:: :      :   : | |  :| :: ||
Db       2 FWGLIMEPNKRYTQVVEKP--FHISQAAM-DISTGDNDPCQVMVVVDGKNFLVCTLQKGK 58

Qy      63 IPQLFCDLVFDKEFELSH-TWGKGSVYFVGYKTPNIEPQGYSEEEEEEEEE-------VP 114
           | |:  || |     :|  | || :|:  ||  |  |     ||| |||||       ||
Db      59 IIQVPLDLYFKSGDSVSFLTNGKCNVHLTGYLDPEFEEDLEDEEEAEEEEEEEEAPPLVP 118

Qy     115 AGNAAK------AVAKPKAKP-----------AEVKPAVDDEE----------------- 140
           | |  |       | |  ||||           ||     |||:
Db     119 AKNKRKLENANDATANKKAKPDKKAGKNSAPAAESDSDDDDEDQLQKFLDGEDIDTDEND 178

Qy     141 -----------DESDSDGMDED----DSDGEDSEEEEPTPKK----PASSKKRANETTPK 181
                      |:|| :  |||    : | || ||||  |||    ||: :    :|: :
Db     179 ESFKMNTSAEGDDSDEEDDDEDEEDEEDDDEDEEEEEAPKKKKKQPAAEQDSTLDTSKE 238

Qy     182 APVSAKKAKVAVTPQKTDEKKKGGKAANQSPK 213
           :   :| :|     ||   |||  : | | |:
Db     239 SVDMSKLSK----SQKRRLKKKLQQQAKQQPQ 266


RESULT   2
FKB4_DROME
ID   FKB4_DROME     STANDARD;      PRT;   357 AA.
AC   P54397; Q9VP88;
DT   01-OCT-1996 (Rel. 34, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   39 kDa FK506-binding nuclear protein (Peptidyl-prolyl cis-trans
DE   isomerase) (PPiase) (EC 5.2.1.8).
GN   FK506-BP1 OR FKBP39 or CG6226.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CANTON-S;
RX   MEDLINE=95278752; PubMed=7538962;
RA   Theopold U., Dal Zotto L., Hultmark D.;
RT   "FKBP39, a Drosophila member of a family of proteins that bind the
RT   immunosuppressive drug FK506.";
RL   Gene 156:247-251(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- FUNCTION: PPIASES ACCELERATE THE FOLDING OF PROTEINS. FKBP506
CC       MAY FUNCTION IN A SIGNAL TRANSDUCTION CASCADE DURING EARLY
CC       DEVELOPMENT.
CC   -!- CATALYTIC ACTIVITY: CIS-TRANS ISOMERIZATION OF PROLINE IMIDIC
CC       PEPTIDE BONDS IN OLIGOPEPTIDES.
CC   -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC   -!- TISSUE SPECIFICITY: UBIQUITOUSLY EXPRESSED, HIGHEST LEVELS IN
CC       OVARY.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED DURING ALL STAGES OF DEVELOPMENT
CC       WITH HIGHEST EXPRESSION IN EARLY EMBRYO.
CC   -!- SIMILARITY: BELONGS TO THE FKBP-TYPE PPIASE FAMILY.
CC   -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-3 IS THE INITIATOR.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z46894; CAA86996.1; -.
DR   EMBL; AE003708; AAF55172.1; ALT_SEQ.
DR   HSSP; P27124; 1ROT.
DR   FlyBase; FBgn0013269; FK506-bp1.
DR   InterPro; IPR001179; FKBP_PPIase.
DR   Pfam; PF00254; FKBP; 1.
DR   PROSITE; PS00453; FKBP_PPIASE_1; FALSE_NEG.
DR   PROSITE; PS00454; FKBP_PPIASE_2; 1.
DR   PROSITE; PS50059; FKBP_PPIASE_3; 1.
KW   Isomerase; Rotamase; Nuclear protein.
FT   DOMAIN       89     99       ASP/GLU-RICH (HIGHLY ACIDIC).
FT   DOMAIN      119    183       ASP/GLU-RICH (HIGHLY ACIDIC).
FT   DOMAIN      186    247       LYS-RICH (BASIC).
FT   DOMAIN      269    357       PPIASE, FKBP-TYPE.
FT   CONFLICT    187    187       A -> R (IN REF. 1).
SQ   SEQUENCE   357 AA;  39343 MW;  EF0AB7831738BB30 CRC64;


  Query Match             15.0%;  Score 191.5;  DB 1;  Length 357;
  Best Local Similarity   25.4%;  Pred. No. 1.4e-05;
  Matches   72;  Conservative   42;  Mismatches  110;  Indels   59;  Gaps   10;

Qy       3 FWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLSTEN 62
           |||: :|  :  : |  :    |:|  :|      ||:   |::    |  :: |: |:
Db       4 FWGLNMKPERKYSQTIIKS--FHISGVAL-----DKGQEAKLYLAAEKQEYIVATV-TKA 55

Qy      63 IPQLFCDLVFDK-EFELSHTWGKGSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGNAAKA 121
           |||:  || | | :  : :| |  ||  :||           |::|::::::   |   :
Db      56 IPQVALDLNFSKGDRIMFYTAGDASVSLLGYL------HDIDSEDDEDDDQMTIENLLNS 109

Qy     122 VAKPKAKPAEVKPAVDD-------EEDESDSD--------------GMDEDDSDGEDSE 159
             |   :| :|     ||       |||| |:|              | :||| | ::
Db     110 KAIKNSKKSE-----DDEDENESGEEDEEDTDDDSQIIEEYESFLENGEEEDDDDVDEDN 164

Qy     160 EEEPTPKKPASSKKRANETTPKAPV-------SAKKA----------KVAVTPQKTDEKK 202
           ||     : |    | |  ||| |       ||||:          : |   ||  ||
```

```
Db    165 EESGEEDEQDSDDSEAEEEQPKAKVAKLSPGASAKKSGKEQNGVAKKEEAKQQQKKKEKP 224

Qy    203 KGGKAANQSPKSASQVSCGSCKKTFNSGNALESHNKAKHAAAK 245
          :  |   :: : | |       :|   |  :      |   ||
Db    225 EAKKEQPKAKEPAKQQPASKDPRTITGGVKIVDQVVGKGEEAK 267


RESULT   3
NPM_XENLA
ID   NPM_XENLA      STANDARD;      PRT;   299 AA.
AC   P07222;
DT   01-APR-1988 (Rel. 07, Created)
DT   01-APR-1988 (Rel. 07, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Nucleophosmin (NPM) (Nucleolar phosphoprotein B23) (Numatrin)
DE   (Nucleolar protein NO38).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=88004394; PubMed=3308448;
RA   Schmidt-Zachmann M.S., Huegle-Doerr B., Franke W.W.;
RT   "A constitutive nucleolar protein identified as a member of the
RT   nucleoplasmin family.";
RL   EMBO J. 6:1881-1890(1987).
CC   -!- FUNCTION: ASSOCIATED WITH NUCLEOLAR RIBONUCLEOPROTEIN STRUCTURES
CC       AND BIND SINGLE-STRANDED NUCLEIC ACIDS. IT MAY FUNCTION IN THE
CC       ASSEMBLY AND/OR TRANSPORT OF RIBOSOME (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC   -!- PTM: PHOSPHORYLATED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE NUCLEOPLASMIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X05496; CAA29046.1; -.
DR   PIR; A29681; A29681.
DR   InterPro; IPR004301; Nucleoplasmin.
DR   Pfam; PF03066; Nucleoplasmin; 1.
KW   Nuclear protein; Phosphorylation; RNA-binding; Repeat.
FT   DOMAIN        1      7       MET-RICH.
FT   DOMAIN      122    137       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      153    158       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   DOMAIN      160    187       ASP/GLU-RICH (HIGHLY ACIDIC).
FT   DOMAIN      189    195       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   DOMAIN      218    242       4 X 3 AA REPEATS OF K-T-P.
FT   REPEAT      218    220
FT   REPEAT      221    223
FT   REPEAT      237    239
FT   REPEAT      240    242
SQ   SEQUENCE   299 AA;  33532 MW;  C16CDF9565090843 CRC64;


  Query Match             13.3%;  Score 170.5;  DB 1;  Length 299;
  Best Local Similarity   31.5%;  Pred. No. 0.00025;
  Matches   56;  Conservative   20;  Mismatches   67;  Indels   35;  Gaps    7;

Qy     83 GKGSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGNAAKAVAKPKAKPAEVKPAVDDEEDE 142
          | | ||  |     :|    |:::|:|| |  |  || :      :|    |  :::||::
Db    107 GSGPVYVSGQHLVALEDLESSDDEDEEHEPSPK-NAKRIAPDSASKVPRKKTRLEEEEED 165

Qy    143 SD-SDGMDEDDSDGEDSEEEEPTP-KKPASSKKRA---------------NETTPKAPVS 185
          ||  |  |||| | :| |||| || ||  |:| :|                : ||| |
Db    166 SDEDDDDDEDDDDEDDDEEEEETPVKKTDSTKSKAAQKLNHNGKASALSTTQKTPKTPEQ 225

Qy    186 AKK-------AKVAVTPQKTDEKK--------KGGKAANQSPKSASQVSCGSCKKTFN 228
          |         |   ||  ::| |         ||        | |: |   :| :| |
Db    226 KGKQDTKPQTPKTPKTPLSSEEIKAKMQTYLEKGNVLPKVEVKFANYVK--NCFRTEN 281


RESULT   4
NUCL_CHICK
ID   NUCL_CHICK      STANDARD;      PRT;   694 AA.
AC   P15771;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Nucleolin (Protein C23).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90206792; PubMed=2320420;
RA   Maridor G., Nigg E.A.;
RT   "cDNA sequences of chicken nucleolin/C23 and NO38/B23, two major
RT   nucleolar proteins.";
RL   Nucleic Acids Res. 18:1286-1286(1990).
RN   [2]
RP   DISCUSSION OF SEQUENCE.
RX   MEDLINE=90304215; PubMed=2114180;
RA   Maridor G., Krek W., Nigg E.A.;
RT   "Structure and developmental expression of chicken nucleolin and
RT   NO38: coordinate expression of two abundant non-ribosomal nucleolar
RT   proteins.";
RL   Biochim. Biophys. Acta 1049:126-133(1990).
RN   [3]
RP   SEQUENCE OF 407-694 FROM N.A.
RX   MEDLINE=89119560; PubMed=2914325;
RA   Borer R.A., Lehner C.F., Eppenberger H.M., Nigg E.A.;
RT   "Major nucleolar proteins shuttle between nucleus and cytoplasm.";
RL   Cell 56:379-390(1989).
CC   -!- FUNCTION: NUCLEOLIN IS THE MAJOR NUCLEOLAR PROTEIN OF GROWING
CC       EUKARYOTIC CELLS. IT IS FOUND ASSOCIATED WITH INTRANUCLEOLAR
CC       CHROMATIN AND PRERIBOSOMAL PARTICLES. IT INDUCES CHROMATIN
CC       DECONDENSATION BY BINDING TO HISTONE H1. IT IS THOUGHT TO PLAY A
CC       ROLE IN PRE-RRNA TRANSCRIPTION AND RIBOSOME ASSEMBLY.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
CC   -!- SIMILARITY: CONTAINS 4 RNA RECOGNITION MOTIFS (RRM).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X17199; CAA35060.1; -.
DR   EMBL; M21791; AAA48983.1; -.
DR   PIR; S08414; DNCHNL.
DR   HSSP; P09651; 1UP1.
DR   InterPro; IPR000504; RRM.
DR   Pfam; PF00076; rrm; 4.
DR   SMART; SM00360; RRM; 4.
DR   PROSITE; PS50102; RRM; 4.
DR   PROSITE; PS00030; RRM_RNP_1; 3.
KW   Nuclear protein; Phosphorylation; Methylation; DNA-binding; Repeat;
KW   RNA-binding.
FT   DOMAIN      124    141       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      170    192       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      217    247       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      281    357       RNA-BINDING (RRM) 1.
FT   DOMAIN      371    445       RNA-BINDING (RRM) 2.
FT   DOMAIN      461    535       RNA-BINDING (RRM) 3.
```

```
FT   DOMAIN      553    628       RNA-BINDING (RRM) 4.
FT   DOMAIN      630    682       ARG/GLY/PHE-RICH.
FT   DOMAIN       55     90       5 X 7 AA TANDEM REPEATS OF X-T-P-X-K-K-X.
FT   REPEAT       55     61       1.
FT   REPEAT       62     68       2.
FT   REPEAT       69     75       3.
FT   REPEAT       76     82       4.
FT   REPEAT       84     90       5.
FT   MOD_RES     116    116       PHOSPHORYLATION (BY SIMILARITY).
FT   MOD_RES     136    136       PHOSPHORYLATION (BY SIMILARITY).
FT   MOD_RES     171    171       PHOSPHORYLATION (BY SIMILARITY).
FT   CONFLICT    419    419       A -> R (IN REF. 3).
FT   CONFLICT    520    520       N -> T (IN REF. 3).
SQ   SEQUENCE   694 AA;  75640 MW;  7996C504BE9459A1 CRC64;


  Query Match           13.3%;  Score 170;  DB 1;  Length 694;
  Best Local Similarity  33.9%;  Pred. No. 0.00065;
  Matches   60;  Conservative   16;  Mismatches   61;  Indels   40;  Gaps    8;

Qy      84 KGSVYFVGYKTPNIEPQGYSEEEEEEE------------EEVPAGNAAKAVAKP---KAK 128
           | |   |  | | : |     ||||||            |:          | ||   ||
Db     151 KKSAAAVPAKKPAVVPAKQESEEEEEEDDEEEDEEDDESEDEAMDTTPAPVKKPTPAKAT 210

Qy     129 PAEVKPAVDDEEDESDSDGMDEDDSDGEDSEEEEPTPKKPAS------SKKRANETTPKA 182
           ||: |   :|||||  | |  ||:| | |: :|||   :||         |: ||:: |:|
Db     211 PAKAKAESEDEEDEEDED-EDEEDEDDEEEDEEESEDEKPVKEAPGKRKKEMANKSAPEA 269

Qy     183 -------PVSAKKAKVA-VTPQKTDE------KKKGGKAANQSPKSASQVSCGSCKK 225
                 | ||     |  :|| | |      |:  ||    |     |:|  || |:
Db     270 KKKKTETPASAFSLFVKNLTPTKDYEELRTAIKEFFGKKNLQ----VSEVRIGSSKR 322


RESULT   5
NPM_CHICK
ID   NPM_CHICK      STANDARD;      PRT;   294 AA.
AC   P16039;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Nucleophosmin (NPM) (Nucleolar phosphoprotein B23) (Numatrin)
DE   (Nucleolar protein NO38).
GN   NPM1.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90206792; PubMed=2320420;
RA   Maridor G., Nigg E.A.;
RT   "cDNA sequences of chicken nucleolin/C23 and NO38/B23, two major
RT   nucleolar proteins.";
RL   Nucleic Acids Res. 18:1286-1286(1990).
RN   [2]
RP   DISCUSSION OF SEQUENCE.
RX   MEDLINE=90304215; PubMed=2114180;
RA   Maridor G., Krek W., Nigg E.A.;
RT   "Structure and developmental expression of chicken nucleolin and
RT   NO38: coordinate expression of two abundant non-ribosomal nucleolar
RT   proteins.";
RL   Biochim. Biophys. Acta 1049:126-133(1990).
CC   -!- FUNCTION: ASSOCIATED WITH NUCLEOLAR RIBONUCLEOPROTEIN STRUCTURES
CC       AND BIND SINGLE-STRANDED NUCLEIC ACIDS. IT MAY FUNCTION IN THE
CC       ASSEMBLY AND/OR TRANSPORT OF RIBOSOME.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR. GENERALLY NUCLEOLAR, BUT IS
CC       TRANSLOCATED TO THE NUCLEOPLASM IN CASE OF SERUM STARVATION OR
CC       TREATMENT WITH ANTICANCER DRUGS (BY SIMILARITY).
CC   -!- PTM: PHOSPHORYLATED.
CC   -!- SIMILARITY: BELONGS TO THE NUCLEOPLASMIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X17200; CAA35061.1; -.
DR   PIR; S08415; DNCHFM.
DR   InterPro; IPR004301; Nucleoplasmin.
DR   Pfam; PF03066; Nucleoplasmin; 1.
KW   Nuclear protein; Phosphorylation; RNA-binding.
FT   DOMAIN        1     11       MET-RICH.
FT   DOMAIN      122    133       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      153    158       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   DOMAIN      161    186       ASP/GLU-RICH (HIGHLY ACIDIC).
FT   DOMAIN      190    196       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ   SEQUENCE   294 AA;  32632 MW;  B1FF89B5F2322DED CRC64;


  Query Match           13.2%;  Score 169;  DB 1;  Length 294;
  Best Local Similarity  34.9%;  Pred. No. 0.0003;
  Matches   51;  Conservative   24;  Mismatches   57;  Indels   14;  Gaps    7;

Qy      76 FELSHTWGKGSVYFVGYKTPNIEPQGYSEEEEEEEEVPAGNAA--KAVAKPKAKPAEVK 133
           | |    | | ||  |     :| :    | |:|||:   |||:  :  :   ||  : |
Db     100 FVLRLKCGSGPVYVSGQHLVALEEE---PESEDEEEDTKIGNASTKRPASGGGAKTPQKK 156

Qy     134 PAV--DDEEDESDSDGMDEDDSDGEDSEEEEPTP-KKPASSKKRANETTPKAPVSAKKAK 190
           | :  |||:|: | |  ::|: | :| ||| || ||||  :: | :   ||  : | :|
Db     157 PKLSEDDEDDDEDEDDDEDDEDDLDDDEEEIKTPMKKPA--REPAGKNMQKAKQNGKDSK 214

Qy     191 VAVTP---QKTDEKKKGGKAANQSPK 213
           : ||    || : ||      ::||
Db     215 PS-TPASKTKTPDSKKDKSLTPKTPK 239


RESULT   6
NUCL_MESAU
ID   NUCL_MESAU     STANDARD;      PRT;   713 AA.
AC   P08199;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Nucleolin (Protein C23).
GN   NCL.
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE OF 35-713 FROM N.A., AND SEQUENCE OF 1-34.
RX   MEDLINE=87175501; PubMed=3470736;
RA   Lapeyre B., Bourbon H., Amalric F.;
RT   "Nucleolin, the major nucleolar protein of growing eukaryotic cells:
RT   an unusual protein structure revealed by the nucleotide sequence.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:1472-1476(1987).
RN   [2]
RP   FUNCTION.
RX   MEDLINE=88312631; PubMed=3409881;
RA   Erard M.S., Belenguer P., Caizergues-Ferrer M., Pantaloni A.,
RA   Amalric F.;
RT   "A major nucleolar protein, nucleolin, induces chromatin
RT   decondensation by binding to histone H1.";
RL   Eur. J. Biochem. 175:525-530(1988).
CC   -!- FUNCTION: NUCLEOLIN IS THE MAJOR NUCLEOLAR PROTEIN OF GROWING
CC       EUKARYOTIC CELLS. IT IS FOUND ASSOCIATED WITH INTRANUCLEOLAR
CC       CHROMATIN AND PRERIBOSOMAL PARTICLES. IT INDUCES CHROMATIN
CC       DECONDENSATION BY BINDING TO HISTONE H1. IT IS THOUGHT TO PLAY A
CC       ROLE IN PRE-RRNA TRANSCRIPTION AND RIBOSOME ASSEMBLY.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
CC   -!- SIMILARITY: CONTAINS 4 RNA RECOGNITION MOTIFS (RRM).
```

```
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M15825; AAA36966.1; -.
DR   PIR; A27441; A27441.
DR   HSSP; P09651; 1HA1.
DR   InterPro; IPR000504; RRM.
DR   Pfam; PF00076; rrm; 4.
DR   SMART; SM00360; RRM; 4.
DR   PROSITE; PS50102; RRM; 4.
DR   PROSITE; PS00030; RRM_RNP_1; 3.
KW   Nuclear protein; Phosphorylation; Methylation; DNA-binding; Repeat;
KW   RNA-binding.
FT   INIT_MET      0      0
FT   DOMAIN      141    169       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      188    213       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      239    271       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      307    383       RNA-BINDING (RRM) 1.
FT   DOMAIN      393    466       RNA-BINDING (RRM) 2.
FT   DOMAIN      485    559       RNA-BINDING (RRM) 3.
FT   DOMAIN      571    646       RNA-BINDING (RRM) 4.
FT   DOMAIN      648    701       ARG/GLY/PHE-RICH.
FT   DOMAIN       56    133       8 X 8 AA TANDEM REPEATS OF X-T-P-X-K-K-
FT                               X-X.
FT   REPEAT       56     63       1.
FT   REPEAT       73     80       2.
FT   REPEAT       81     88       3.
FT   REPEAT       89     96       4.
FT   REPEAT       97    102       5 (INCOMPLETE).
FT   REPEAT      103    110       6.
FT   REPEAT      118    125       7.
FT   REPEAT      126    133       8.
FT   REPEAT      217    221       X-T-P-X-K-K-X-X MOTIF.
FT   MOD_RES     143    143       PHOSPHORYLATION.
FT   MOD_RES     156    156       PHOSPHORYLATION.
FT   MOD_RES     187    187       PHOSPHORYLATION.
FT   MOD_RES     655    655       METHYLATION (DI-).
FT   MOD_RES     659    659       METHYLATION (DI-).
FT   MOD_RES     665    665       METHYLATION (DI-).
FT   MOD_RES     669    669       METHYLATION (DI-).
FT   MOD_RES     673    673       METHYLATION (DI-).
FT   MOD_RES     679    679       METHYLATION (DI-).
FT   MOD_RES     681    681       METHYLATION (DI-).
FT   MOD_RES     687    687       METHYLATION (DI-).
FT   MOD_RES     691    691       METHYLATION (DI-).
FT   MOD_RES     694    694       METHYLATION (DI-).
SQ   SEQUENCE   713 AA;  76997 MW;  79DDCF724CED7DB4 CRC64;


  Query Match             13.0%;  Score 166.5;  DB 1;  Length 713;
  Best Local Similarity   33.8%;  Pred. No. 0.0011;
  Matches   48;  Conservative   20;  Mismatches   55;  Indels   19;  Gaps    4;

Qy      73 DKEFELSHTWGK-GSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGNAAKAV----AKPKA 127
           : |||     || | |       : : : :   |||||:|||        :|:     || |
Db     165 EDEFEPPVVKGKQGKVAAAAPASEDEDEEEDEEEEEEDEEEEDDSEEEEAMEITPAKGKK 224

Qy     128 KPAEVKPAV----------DDEEDESDSDGMDEDDSDGEDSEEEEPTPKKPASSKKRANE 177
             ||:| |            |:|||| : :  :|:: : |: ||||  | |||  |::
Db     225 APAKVVPVKAKNVAEEDDDDEEEDEDEEEDEEEEEDEEEEEEEEEEEPVKPAPGKRKKEM 284

Qy     178 TTPKAPVSAKKAKV----AVTP 195
           |  |    ||| ||    : ||
Db     285 TKQKEVPEAKKQKVEGSESTTP 306


RESULT   7
```

```
NPM_HUMAN
ID   NPM_HUMAN      STANDARD;      PRT;   294 AA.
AC   P06748; P08693; Q12826; Q14115; Q13440; Q13441;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Nucleophosmin (NPM) (Nucleolar phosphoprotein B23) (Numatrin)
DE   (Nucleolar protein NO38).
GN   NPM1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89229105; PubMed=2713355;
RA   Chan W.-Y., Liu Q.R., Borjigin J., Busch H., Rennert O.M., Tease L.A.,
RA   Chan P.-K.;
RT   "Characterization of the cDNA encoding human nucleophosmin and
RT   studies of its role in normal and abnormal growth.";
RL   Biochemistry 28:1033-1039(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89374300; PubMed=2775293;
RA   Li X., McNeilage L.J., Whittingham S.;
RT   "The nucleotide sequence of a human cDNA encoding the highly
RT   conserved nucleolar phosphoprotein B23.";
RL   Biochem. Biophys. Res. Commun. 163:72-78(1989).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90026373; PubMed=2478125;
RA   Zhang X.T., Thomis D.C., Samuel C.E.;
RT   "Isolation and characterization of a molecular cDNA clone of a human
RT   mRNA from interferon-treated cells encoding nucleolar protein B23,
RT   numatrin.";
RL   Biochem. Biophys. Res. Commun. 164:176-184(1989).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97248692; PubMed=9092633;
RA   Chan P.-K., Chan F.Y., Morris S.W., Xie Z.;
RT   "Isolation and characterization of the human nucleophosmin/B23 (NPM)
RT   gene: identification of the YY1 binding site at the 5' enhancer
RT   region.";
RL   Nucleic Acids Res. 25:1225-1232(1997).
RN   [5]
RP   SEQUENCE OF 15-294 FROM N.A.
RX   MEDLINE=90098787; PubMed=2602120;
RA   Hale T.K., Mansfield B.C.;
RT   "Nucleotide sequence of a cDNA clone representing a third allele of
RT   human protein B23.";
RL   Nucleic Acids Res. 17:10112-10112(1989).
RN   [6]
RP   SEQUENCE OF 213-294 FROM N.A., AND SEQUENCE OF 227-294.
RX   MEDLINE=87033628; PubMed=2429957;
RA   Chan P.-K., Chan W.-Y., Yung B.Y.M., Cook R.G., Aldrich M.B.,
RA   Ku D., Goldknopf I.L., Busch H.;
RT   "Amino acid sequence of a specific antigenic peptide of protein B23.";
RL   J. Biol. Chem. 261:14335-14341(1986).
RN   [7]
RP   SEQUENCE OF 1-117 FROM N.A.
RC   TISSUE=Lymphoma;
RX   MEDLINE=94167588; PubMed=8122112;
RA   Morris S.W., Kirstein M.N., Valentine M.B., Dittmer K.G.,
RA   Shapiro D.N., Saltman D.L., Look A.T.;
RT   "Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in
RT   non-Hodgkin's lymphoma.";
RL   Science 263:1281-1284(1994).
RN   [8]
RP   SEQUENCE OF 1-117 FROM N.A.
RX   MEDLINE=96210614; PubMed=8633037;
RA   Fujimoto J., Shiota M., Iwahara T., Seki N., Satoh H., Mori S.,
RA   Yamamoto T.;
RT   "Characterization of the transforming activity of p80, a
```

```
RT   hyperphosphorylated protein in a Ki-1 lymphoma cell line with
RT   chromosomal translocation t(2;5).";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:4181-4186(1996).
RN   [9]
RP   SEQUENCE OF 1-133 FROM N.A.
RC   TISSUE=Bone marrow;
RX   MEDLINE=96151966; PubMed=8562957;
RA   Redner R.L., Rush E.A., Faas S., Rudert W.A., Corey S.J.;
RT   "The t(5;17) variant of acute promyelocytic leukemia expresses a
RT   nucleophosmin-retinoic acid receptor fusion.";
RL   Blood 87:882-886(1996).
RN   [10]
RP   SEQUENCE OF 115-134.
RX   MEDLINE=86111864; PubMed=3944116;
RA   Chan P.-K., Aldrich M.B., Cook R.G., Busch H.;
RT   "Amino acid sequence of protein B23 phosphorylation site.";
RL   J. Biol. Chem. 261:1868-1872(1986).
RN   [11]
RP   SEQUENCE OF 33-54.
RC   TISSUE=Colon carcinoma;
RX   MEDLINE=97295306; PubMed=9150948;
RA   Ji H., Reid G.E., Moritz R.L., Eddes J.S., Burgess A.W., Simpson R.J.;
RT   "A two-dimensional gel database of human colon carcinoma proteins.";
RL   Electrophoresis 18:605-613(1997).
RN   [12]
RP   CHROMOSOMAL TRANSLOCATION.
RX   MEDLINE=96152893; PubMed=8570204;
RA   Yoneda-Kato N., Look A.T., Kirstein M.N., Valentine M.B.,
RA   Raimondi S.C., Cohen K.J., Carroll A.J., Morris S.W.;
RT   "The t(3;5)(q25.1;q34) of myelodysplastic syndrome and acute myeloid
RT   leukemia produces a novel fusion gene, NPM-MLF1.";
RL   Oncogene 12:265-275(1996).
CC   -!- FUNCTION: ASSOCIATED WITH NUCLEOLAR RIBONUCLEOPROTEIN STRUCTURES
CC       AND BIND SINGLE-STRANDED NUCLEIC ACIDS. IT MAY FUNCTION IN THE
CC       ASSEMBLY AND/OR TRANSPORT OF RIBOSOME.
CC   -!- SUBUNIT: DISULFIDE-LINKED DIMERS UNDER CERTAIN CONDITIONS.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR, GENERALLY NUCLEOLAR, BUT IS
CC       TRANSLOCATED TO THE NUCLEOPLASM IN CASE OF SERUM STARVATION OR
CC       TREATMENT WITH ANTICANCER DRUGS.
CC   -!- PTM: PHOSPHORYLATED.
CC   -!- DISEASE: A FORM OF NON-HODGKIN'S LYMPHOMA IS CHARACTERIZED BY A
CC       CHROMOSOMAL TRANSLOCATION T(2;5)(P23;Q35) THAT INVOLVES NPM AND
CC       ALK.
CC   -!- DISEASE: A FORM OF ACUTE PROMYELOCYTIC LEUKEMIA IS CHARACTERIZED
CC       BY A CHROMOSOMAL TRANSLOCATION T(5;17)(Q32;Q11) THAT INVOLVES NPM
CC       AND RARA.
CC   -!- DISEASE: PARTICIPATES IN A T(3;5)(Q25.1;Q34) CHROMOSOMAL
CC       TRANSLOCATION. THE PRODUCT, A NPM-MLF1 FUSION PROTEIN, IS
CC       RESPONSIBLE FOR MYELODYSPLASTIC SYNDROME (MDS) WHICH PROGRESSES TO
CC       ACUTE MYELOID LEUKEMIA (AML).
CC   -!- SIMILARITY: BELONGS TO THE NUCLEOPLASMIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M23613; AAA36380.1; -.
DR   EMBL; M28699; AAA58386.1; -.
DR   EMBL; M26697; AAA36385.1; -.
DR   EMBL; U89321; AAB94739.1; -.
DR   EMBL; U89309; AAB94739.1; JOINED.
DR   EMBL; U89310; AAB94739.1; JOINED.
DR   EMBL; U89311; AAB94739.1; JOINED.
DR   EMBL; U89313; AAB94739.1; JOINED.
DR   EMBL; U89314; AAB94739.1; JOINED.
DR   EMBL; U89317; AAB94739.1; JOINED.
DR   EMBL; U89319; AAB94739.1; JOINED.
DR   EMBL; X16934; CAA34809.1; -.
DR   EMBL; J02590; AAA36473.1; -.
DR   EMBL; M31004; AAA36474.1; ALT_SEQ.
DR   EMBL; U04946; AAA58698.1; ALT_TERM.
DR   EMBL; D45915; BAA08343.1; ALT_TERM.
DR   EMBL; U41742; AAB00112.1; ALT_TERM.
DR   EMBL; U41743; AAB00113.1; ALT_TERM.
DR   PIR; A30137; A30137.
DR   PIR; A32915; A32915.
DR   PIR; A26080; A26080.
DR   PIR; B26080; B26080.
DR   PIR; S06926; S06926.
DR   PIR; A33423; A33423.
DR   SWISS-2DPAGE; P06748; HUMAN.
DR   MIM; 164040; -.
DR   InterPro; IPR004301; Nucleoplasmin.
DR   Pfam; PF03066; Nucleoplasmin; 1.
KW   Nuclear protein; Phosphorylation; RNA-binding; Proto-oncogene;
KW   Chromosomal translocation.
FT   DOMAIN        1      9       MET-RICH.
FT   DOMAIN      120    132       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      152    157       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   DOMAIN      161    188       ASP/GLU-RICH (HIGHLY ACIDIC).
FT   DOMAIN      191    197       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   MOD_RES     125    125       PHOSPHORYLATION.
FT   MOD_RES     199    199       PHOSPHORYLATION.
FT   SITE        175    176       BREAKPOINT FOR TRANSLOCATION TO FORM NPM-
FT                                MLF1.
FT   CONFLICT    129    129       E -> D (IN REF. 10).
FT   CONFLICT    213    213       D -> P (IN REF. 6).
FT   CONFLICT    216    216       P -> S (IN REF. 6).
FT   CONFLICT    219    221       TPR -> SSS (IN REF. 6).
SQ   SEQUENCE   294 AA;  32575 MW;  620BC7BA2E4A0054 CRC64;


  Query Match             12.1%;  Score 155;  DB 1;  Length 294;
  Best Local Similarity   28.3%;  Pred. No. 0.0023;
  Matches   43;  Conservative   25;  Mismatches   56;  Indels   28;  Gaps    5;

Qy      83 GKGSVYFVGYKTPNIEPQGYSEEEEEEE---------EEVPAGNAAKAVAKPKAKPAEVK 133
           | | |:  |     :|    ||:||||:           | |      :|   |  ::
Db     105 GSGPVHISGQHLVAVEEDAESEDEEEEDVKLLSISGKRSAPGGG-----SKVPQKKVKLA 159

Qy     134 PAVDDEEDESDSDGMDEDDSDGEDSEEEEPTPKK------PASSKKRANETTPKAPVSAK 187
             ||::|: : |  |:|| | :| | ||  | |      || : :::|:       : |
Db     160 ADEDDDDDDEEDDDEDDDDDDFDDEEAEEKAPVKKSIRDTPAKNAQKSNQ-------NGK 212

Qy     188 KAKVAVTPQ-KTDEKKKGGKAANQSPKSASQV 218
           :| : ||: |  |  |  |  :   ::||  | |
Db     213 DSKPSSTPRSKGQQESFKKQEKTPKTPKGPSSV 244


RESULT   8
MP62_LYTPI
ID   MP62_LYTPI     STANDARD;      PRT;   411 AA.
AC   P91753;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Mitotic apparatus protein P62.
OS   Lytechinus pictus (Painted sea urchin).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Echinacea; Temnopleuroida; Toxopneustidae;
OC   Lytechinus.
OX   NCBI_TaxID=7653;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97166213; PubMed=9013612;
RA   Ye X., Sloboda R.D.;
RT   "Molecular characterization of p62, a mitotic apparatus protein
RT   required for mitotic progression.";
RL   J. Biol. Chem. 272:3606-3614(1997).
CC   -!- FUNCTION: REQUIRED FOR MITOTIC PROGRESSION. BINDS TO CHROMATIN.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- PTM: PHOSPHORYLATED BY CAM-KINASE II IN VITRO.
```

```
CC   -!- SIMILARITY: BELONGS TO THE NUCLEOPLASMIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U76750; AAB47481.1; -.
DR   InterPro; IPR004301; Nucleoplasmin.
DR   Pfam; PF03066; Nucleoplasmin; 1.
KW   Nuclear protein; Phosphorylation; Mitosis.
FT   DOMAIN      147    156       POLY-GLU.
FT   DOMAIN      192    200       POLY-GLU.
FT   DOMAIN      262    267       POLY-GLU.
FT   DOMAIN      272    284       POLY-GLU.
FT   DOMAIN      293    300       POLY-ASP.
FT   DOMAIN      313    319       POLY-GLU.
SQ   SEQUENCE   411 AA;  46385 MW;  D2DE04293C362254 CRC64;


  Query Match           12.0%;  Score 153.5;  DB 1;  Length 411;
  Best Local Similarity  26.8%;  Pred. No. 0.0041;
  Matches   73;  Conservative   30;  Mismatches   90;  Indels   79;  Gaps   13;

Qy       3 FWGIEVKSGKPV-TVTPEEGIL--------------IHVSQASLG-------------EC 34
           |||  :   | :    ||   |                : : || ||             |
Db       6 FWGATLSKDKKIFKWDPESDFLDDEDDDEEDSISHFLFLKQAVLGVNAKDDDRNVIEVET 65

Qy      35 KNKKCEFV--PL-----------HVKVGNQ-----NLVLGT----LSTENIPQLFCDLVF 72
             |  || |  ||           :: :| |       | ||:    || ::   |  |  |
Db      66 INFDGETVIQPLLSLRLGLNESTNLDIGLQPPVTFKLALGSGPVYLSGQHALDLQEDEEF 125

Qy      73 DKEFELSHTWGKGSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGNAAKAVAKPKAKPAEV 132
             |:||          | || :    | :|   ||:|||| |   : | : |   | | |
Db     126 GKDFE------GAEAYEVGDEDLEDEDEG---EEDEEEEETPKKGSPKRIVK---KIAAV 173

Qy     133 KPAVDDEEDESDSDGMDEDDSDGEDSEEEE---------PTPKKPASSKKRA-----NET 178
           |  :  : || |    |||| : |: ||||         | |     :|| |       :
Db     174 KGRMKGKGDELDE---DEDDDEEEEEEEEEIQTAKGKKRPAPSAKGPAKKLAKVDKDGTS 230

Qy     179 TPKAPVSAKKAKVAVTPQKTDEKKKGGKAANQ 210
             | |  : :   |:   : ||: :  |  ::
Db     231 KRKVPNGSVENGHAIDDDEDDEEDEDYKVGDE 262


RESULT   9
NUCL_MOUSE
ID   NUCL_MOUSE     STANDARD;      PRT;   706 AA.
AC   P09405; Q61991;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Nucleolin (Protein C23).
GN   NCL OR NUC.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C;
RX   MEDLINE=88316930; PubMed=3137346;
RA   Bourbon H., Lapeyre B., Amalric F.;
RT   "Structure of the mouse nucleolin gene. The complete sequence reveals
RT   that each RNA binding domain is encoded by two independent exons.";
RL   J. Mol. Biol. 200:627-638(1988).
RN   [2]
RP   SEQUENCE OF 1-44 FROM N.A.
RX   MEDLINE=89121496; PubMed=2906027;
RA   Bourbon H.M., Prudhomme M., Amalric F.;
RT   "Sequence and structure of the nucleolin promoter in rodents:
RT   characterization of a strikingly conserved CpG island.";
RL   Gene 68:73-84(1988).
RN   [3]
RP   SEQUENCE OF 1-23.
RX   MEDLINE=91317840; PubMed=1860869;
RA   Pasternack M.S., Bleier K.J., McInerney T.N.;
RT   "Granzyme A binding to target cell proteins. Granzyme A binds to and
RT   cleaves nucleolin in vitro.";
RL   J. Biol. Chem. 266:14703-14708(1991).
CC   -!- FUNCTION: NUCLEOLIN IS THE MAJOR NUCLEOLAR PROTEIN OF GROWING
CC       EUKARYOTIC CELLS. IT IS FOUND ASSOCIATED WITH INTRANUCLEOLAR
CC       CHROMATIN AND PRERIBOSOMAL PARTICLES. IT INDUCES CHROMATIN
CC       DECONDENSATION BY BINDING TO HISTONE H1. IT IS THOUGHT TO PLAY A
CC       ROLE IN PRE-RRNA TRANSCRIPTION AND RIBOSOME ASSEMBLY.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
CC   -!- SIMILARITY: CONTAINS 4 RNA RECOGNITION MOTIFS (RRM).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X07699; CAA30538.1; -.
DR   EMBL; M22089; AAA39841.1; -.
DR   PIR; A29958; DNMS.
DR   HSSP; P09651; 1HA1.
DR   MGD; MGI:97286; Ncl.
DR   InterPro; IPR000504; RRM.
DR   Pfam; PF00076; rrm; 4.
DR   SMART; SM00360; RRM; 4.
DR   PROSITE; PS50102; RRM; 4.
DR   PROSITE; PS00030; RRM_RNP_1; 3.
KW   Nuclear protein; Phosphorylation; Methylation; DNA-binding; Repeat;
KW   RNA-binding.
FT   INIT_MET      0      0
FT   DOMAIN      142    169       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      189    214       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      240    272       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      308    384       RNA-BINDING (RRM) 1.
FT   DOMAIN      394    467       RNA-BINDING (RRM) 2.
FT   DOMAIN      486    560       RNA-BINDING (RRM) 3.
FT   DOMAIN      568    643       RNA-BINDING (RRM) 4.
FT   DOMAIN      645    696       ARG/GLY/PHE-RICH.
FT   MOD_RES     144    144       PHOSPHORYLATION (BY SIMILARITY).
FT   MOD_RES     156    156       PHOSPHORYLATION (BY SIMILARITY).
FT   MOD_RES     188    188       PHOSPHORYLATION (BY SIMILARITY).
FT   DOMAIN       57    134       8 X 8 AA TANDEM REPEATS OF X-T-P-X-K-K-
FT                               X-X.
FT   REPEAT       57     64       1.
FT   REPEAT       74     81       2.
FT   REPEAT       82     89       3.
FT   REPEAT       90     97       4.
FT   REPEAT       98    103       5 (INCOMPLETE).
FT   REPEAT      104    111       6.
FT   REPEAT      119    126       7.
FT   REPEAT      127    134       8.
SQ   SEQUENCE   706 AA;  76592 MW;  89505EE39C89F832 CRC64;


  Query Match           12.0%;  Score 153.5;  DB 1;  Length 706;
  Best Local Similarity  24.1%;  Pred. No. 0.0072;
  Matches   61;  Conservative   31;  Mismatches   94;  Indels   67;  Gaps    6;

Qy       6 IEVKSGKPVTVTPEEGILI----------------------------HVSQASLGECK 35
           |  | ||  | || : :::                                ::: | :
Db      47 IPQKKGKKATTTPAKKVVVSQTKKAAVPTPAKKAAVTPGKKAVATPAKKNITPAKVIPTP 106

Qy      36 NKKG-----EFVPLHVKVGNQNLVLGTLSTENIPQLFCDLVFDKEFELSHTWGKGSVYFV 90
              |||       ||   | |      |  : :|  :   |   |:| |      :
```

```
Db    107 GKKGAAQAKALVPTPGKKGAATPAKGAKNGKNAKKEDSDEDEDEEDEDDSDEDEDDEEED 166

Qy     91 GYKTPNI----------------EPQGYSEEEEEEEEEVPAGNAAKAV-----AKPKAK 128
          :: | :                |     |::||:::|    :: : |     || |
Db    167 EFEPPIVKGVKPAKAAPAAPASEDEEDDEDEDDEEDDDEEEEDDSEEEVMEITTAKGKKT 226

Qy    129 PAEVKPAV---------DDEEDESDSDGMD-EDDSDGEDSEEEEPTPKKPASSKKRANET 178
          ||:| |           |:|||| | |  | |:| : :| ||||  | | |  |::   |
Db    227 PAKVVPMKAKSVAEEEDDEEEDEDDEDEDDEEEDDEDDDEEEEEEEPVKAAPGKRKKEMT 286

Qy    179 TPKAPVSAKKAKV 191
            |    ||| ||
Db    287 KQKEAPEAKKQKV 299


RESULT 10
NUCL_HUMAN
ID   NUCL_HUMAN      STANDARD;      PRT;   706 AA.
AC   P19338;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Nucleolin (Protein C23).
GN   NCL.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89290043; PubMed=2737305;
RA   Srivastava M., Fleming P.J., Pollard H.B., Burns A.L.;
RT   "Cloning and sequencing of the human nucleolin cDNA.";
RL   FEBS Lett. 250:99-105(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90368666; PubMed=2394707;
RA   Srivastava M., McBride O.W., Fleming P.J., Pollard H.B., Burns A.L.;
RT   "Genomic organization and chromosomal localization of the human
RT   nucleolin gene.";
RL   J. Biol. Chem. 265:14922-14931(1990).
CC   -!- FUNCTION: NUCLEOLIN IS THE MAJOR NUCLEOLAR PROTEIN OF GROWING
CC       EUKARYOTIC CELLS. IT IS FOUND ASSOCIATED WITH INTRANUCLEOLAR
CC       CHROMATIN AND PRERIBOSOMAL PARTICLES. IT INDUCES CHROMATIN
CC       DECONDENSATION BY BINDING TO HISTONE H1. IT IS THOUGHT TO PLAY A
CC       ROLE IN PRE-RRNA TRANSCRIPTION AND RIBOSOME ASSEMBLY.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
CC   -!- SIMILARITY: CONTAINS 4 RNA RECOGNITION MOTIFS (RRM).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M60858; AAA59954.1; -.
DR   PIR; A35804; A35804.
DR   PIR; S04631; S04631.
DR   HSSP; P09651; 1HA1.
DR   Aarhus/Ghent-2DPAGE; 1210; NEPHGE.
DR   MIM; 164035; -.
DR   InterPro; IPR000504; RRM.
DR   Pfam; PF00076; rrm; 4.
DR   SMART; SM00360; RRM; 4.
DR   PROSITE; PS50102; RRM; 4.
DR   PROSITE; PS00030; RRM_RNP_1; 3.
KW   Nuclear protein; Phosphorylation; Methylation; DNA-binding; Repeat;
KW   RNA-binding.
FT   INIT_MET      0      0
FT   DOMAIN      142    170       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      184    208       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      233    270       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      306    382       RNA-BINDING (RRM) 1.
FT   DOMAIN      392    465       RNA-BINDING (RRM) 2.
FT   DOMAIN      485    559       RNA-BINDING (RRM) 3.
FT   DOMAIN      571    643       RNA-BINDING (RRM) 4.
FT   DOMAIN      645    694       ARG/GLY/PHE-RICH.
FT   DOMAIN       57    134       8 X 8 AA TANDEM REPEATS OF X-T-P-X-K-K-
FT                               X-X.
FT   REPEAT       57     64       1.
FT   REPEAT       74     81       2.
FT   REPEAT       82     89       3.
FT   REPEAT       90     97       4.
FT   REPEAT       98    103       5 (INCOMPLETE).
FT   REPEAT      104    111       6.
FT   REPEAT      119    126       7.
FT   REPEAT      127    134       8.
FT   MOD_RES     144    144       PHOSPHORYLATION (BY SIMILARITY).
FT   MOD_RES     152    152       PHOSPHORYLATION (BY SIMILARITY).
FT   MOD_RES     183    183       PHOSPHORYLATION (BY SIMILARITY).
SQ   SEQUENCE   706 AA;  76213 MW;  85A2F2CA22EA03DB CRC64;


  Query Match             12.0%;  Score 153;  DB 1;  Length 706;
  Best Local Similarity   37.1%;  Pred. No. 0.0077;
  Matches   39;  Conservative    9;  Mismatches   47;  Indels   10;  Gaps    2;

Qy    104 EEEEEEEEEVPAGNAAKAVAKPKAKPAEVKPAVDDE------EDESDSDGMDEDDSDGED 157
          :||:: |||      ||      |  | : |   :||      ||| | |  |::| | ||
Db    200 DEEDDSEEEAMETTPAKGKKAAKVVPVKAKNVAEDEDEEEDDEDEDDDDDEDDEDDDED 259

Qy    158 ----SEEEEPTPKKPASSKKRANETTPKAPVSAKKAKVAVTPQKT 198
                |||| | | |  |::      ||   ||| ||  |   |
Db    260 DEEEEEEEEEEPVKEAPGKRKKEMAKQKAAPEAKKQKVEGTEPTT 304


RESULT 11
NUCL_RAT
ID   NUCL_RAT        STANDARD;      PRT;   712 AA.
AC   P13383;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Nucleolin (Protein C23).
GN   NCL OR NUC.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90269607; PubMed=2347493;
RA   Bourbon H.-M., Amalric F.;
RT   "Nucleolin gene organization in rodents: highly conserved sequences
RT   within three of the 13 introns.";
RL   Gene 88:187-196(1990).
RN   [2]
RP   SEQUENCE OF 1-44 FROM N.A.
RX   MEDLINE=89121496; PubMed=2906027;
RA   Bourbon H.-M., Prudhomme M., Amalric F.;
RT   "Sequence and structure of the nucleolin promoter in rodents:
RT   characterization of a strikingly conserved CpG island.";
RL   Gene 68:73-84(1988).
CC   -!- FUNCTION: NUCLEOLIN IS THE MAJOR NUCLEOLAR PROTEIN OF GROWING
CC       EUKARYOTIC CELLS. IT IS FOUND ASSOCIATED WITH INTRANUCLEOLAR
CC       CHROMATIN AND PRERIBOSOMAL PARTICLES. IT INDUCES CHROMATIN
CC       DECONDENSATION BY BINDING TO HISTONE H1. IT IS THOUGHT TO PLAY A
CC       ROLE IN PRE-RRNA TRANSCRIPTION AND RIBOSOME ASSEMBLY.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
CC   -!- SIMILARITY: CONTAINS 4 RNA RECOGNITION MOTIFS (RRM).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; M55022; AAA41732.1; -.
DR   EMBL; M55015; AAA41732.1; JOINED.
DR   EMBL; M55017; AAA41732.1; JOINED.
DR   EMBL; M55020; AAA41732.1; JOINED.
DR   EMBL; M22090; AAA41733.1; -.
DR   PIR; JH0148; JH0148.
DR   HSSP; P09651; 1HA1.
DR   InterPro; IPR000504; RRM.
DR   Pfam; PF00076; rrm; 4.
DR   SMART; SM00360; RRM; 4.
DR   PROSITE; PS50102; RRM; 4.
DR   PROSITE; PS00030; RRM_RNP_1; 3.
KW   Nuclear protein; Phosphorylation; Methylation; DNA-binding; Repeat;
KW   RNA-binding.
FT   INIT_MET      0      0
FT   DOMAIN      142    167       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      187    215       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      241    274       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      310    386       RNA-BINDING (RRM) 1.
FT   DOMAIN      396    469       RNA-BINDING (RRM) 2.
FT   DOMAIN      488    562       RNA-BINDING (RRM) 3.
FT   DOMAIN      574    649       RNA-BINDING (RRM) 4.
FT   DOMAIN      651    702       ARG/GLY/PHE-RICH.
FT   DOMAIN       57    134       8 X 8 AA TANDEM REPEATS OF X-T-P-X-K-K-
FT                               X-X.
FT   REPEAT       57     64       1.
FT   REPEAT       74     81       2.
FT   REPEAT       82     89       3.
FT   REPEAT       90     97       4.
FT   REPEAT       98    103       5 (INCOMPLETE).
FT   REPEAT      104    111       6.
FT   REPEAT      119    126       7.
FT   REPEAT      127    134       8.
SQ   SEQUENCE   712 AA;  77016 MW;  68774A214E550F90 CRC64;
```

```
  Query Match             11.8%;  Score 151.5;  DB 1;  Length 712;
  Best Local Similarity   24.6%;  Pred. No. 0.0096;
  Matches   65;  Conservative   24;  Mismatches   88;  Indels   87;  Gaps    8;

Qy       6 IEVKSGKPVTVTPEEGILIH------------------------------VSQASLGECK 35
           |  | ||  | || : :::                              |: | :
Db      47 IPQKKGKKATTTPAKKVVVSQTKKAAVPTPAKKAAVTPGKKAAATPAKKAVTPAKVVPTP 106

Qy      36 NKKG-----EFVPLHVKVGNQNLVLGTLSTENIPQLFCDLVFDK--------------EF 76
             |||       ||   | |      |  : :|  :   |   |:               ||
Db     107 GKKGAAQAKALVPTPGKKGAVTPAKGAKNGKNAKKEDSDEDEDEEDEDDSDEDEDEEDEF 166

Qy      77 ELSHTWGKGSVYFVGYKTPNIEPQGYSEEEEEE------------EEEVPAGNAAKAV-- 122
           |            | |     |   : |:|:|            |||    :: : |
Db     167 E--------PPVVKGVKPAKAAPAAPASEDEDEEDDDDEDDDDDDEEEEEEDDSEEEVME 218

Qy     123 ---AKPKAKPAEVKPA------------VDDEEDESDSDGMDEDDSDGEDSEEEEPTPKK 167
              || |  ||:| |              |||::| | |  ||:| : || ||||  | |
Db     219 ITPAKGKKTPAKVVPVKAKSVAEEEEDDEDDEDEEEDEDEEDEEDDEDEDEEEEE-EPVK 277

Qy     168 PASSKKRANETTPKAPVSAKKAKV 191
           |  |::   |  |    ||| |:
Db     278 AAPGKRKKEMTKQKEAPEAKKQKI 301
```

```
RESULT  12
HS32_DICDI
ID   HS32_DICDI     STANDARD;      PRT;   281 AA.
AC   P54658;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   32 kDa heat shock protein (4-1 protein).
GN   HSPC OR HSP32.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96001265; PubMed=7557471;
RA   de Maria A.C., Gomes S.L., Juliania M.H., Mazzarella R., Klein C.;
RT   "Cloning of a cDNA encoding a novel heat-shock protein from
RT   Dictyostelium discoideum.";
RL   Gene 163:163-164(1995).
CC   -!- DEVELOPMENTAL STAGE: PRESENT AT HIGH LEVELS IN GROWING CELLS BUT
CC       DECREASES DRAMATICALLY DURING THE EARLY HOURS OF DEVELOPMENT.
CC   -!- INDUCTION: BY HEAT SHOCK.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U40211; AAC47710.1; -.
DR   EMBL; L39778; AAA99510.1; -.
DR   DictyDb; DD01076; hspC.
KW   Heat shock.
FT   DOMAIN      251    257       POLY-GLN.
FT   DOMAIN      262    265       POLY-ASN.
SQ   SEQUENCE   281 AA;  31481 MW;  2C64F57C847D7CA4 CRC64;
```

```
  Query Match             11.8%;  Score 150.5;  DB 1;  Length 281;
  Best Local Similarity   23.4%;  Pred. No. 0.0042;
  Matches   66;  Conservative   47;  Mismatches   94;  Indels   75;  Gaps   11;

Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGE-FVPLHVKVGNQN------- 52
           |:|:|  |   :|| :  :|| : |:::| :      ||: ::   : :  ::
Db       2 MQFFGTIVTKEEPVNLELDEGDIFHLTKAIIHPKSQGKGKVYLTAVISLMEEDEMEEDDV 61

Qy      53 -------------LVLGTLSTENIPQLFCDLVFD----KEFELSHTWGKGSVYFVGYKTP 95
                       : :| |    | |:  :| ::      |||    | | |  :
Db      62 DDEEESPREDIVEIPIGILEAGKIDQIDLNLHYNFGQIVRFELQAENGAGYVVALSGSVI 121

Qy      96 NIEPQGYSEEEEEEEEEVPAGNAAKAVAKPKAKPAEVKPAVDDEEDESDSDGMDEDDSDG 155
            :|  |  :|: ::|  :                        |||| :|| :   : | |
Db     122 TMEQGGCDDEDCDDEHCI-------------------NHEDDEEIDSDEE-FGDSDQDE 160

Qy     156 EDSEEEE------PTPKK---------PASSKKRANETTPKAPVSAKKAKVAVTPQKTDE 200
           |||::||       |  ||         | | |::  |  ||     || :|    | |
Db     161 EDSDDEEIPQLIAPATKKGKITEISEVPESKKEKTPE--PKKVPEPKKEQV---KQPTQP 215

Qy     201 KKKGGKAANQSPKSASQVSCGSCKKTFNSGNALESHNKAKHA 242
           ::|  ||| | |: |:         |   :  |   :|::|:|
Db     216 QQK--KAAAQQPEKAN-------NKPAAASPAKPQNNQSKNA 248
```

```
RESULT  13
NFM_RABIT
ID   NFM_RABIT      STANDARD;      PRT;   644 AA.
AC   P54938;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Neurofilament triplet M protein (160 kDa neurofilament protein)
DE   (Neurofilament medium polypeptide) (NF-M) (Fragment).
GN   NEF3 OR NEFM OR NFM.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   TISSUE=Heart;
RX   MEDLINE=97055255; PubMed=8899542;
RA   Vitadello M., Vettore S., Lamar E., Chien K.R., Gorza L.;
RT   "Neurofilament M mRNA is expressed in conduction system myocytes of
RT   the developing and adult rabbit heart.";
RL   J. Mol. Cell. Cardiol. 28:1833-1844(1996).
CC   -!- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
CC       AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
CC   -!- PTM: THERE ARE A NUMBER OF REPEATS OF THE TRIPEPTIDE K-S-P, NFM IS
CC       PHOSPHORYLATED ON A NUMBER OF THE SERINES IN THIS MOTIF. IT IS
CC       THOUGHT THAT PHOSPHORYLATION OF NFH RESULTS IN THE FORMATION OF
CC       INTERFILAMENT CROSS BRIDGES THAT ARE IMPORTANT IN THE MAINTENANCE
CC       OF AXONAL CALIBER.
CC   -!- PTM: PHOSPHORYLATION SEEMS TO PLAY A MAJOR ROLE IN THE FUNCTIONING
CC       OF THE LARGER NEUROFILAMENT POLYPEPTIDES (NF-M AND NF-H), THE
CC       LEVELS OF PHOSPHORYLATION BEING ALTERED DEVELOPMENTALLY AND
CC       COINCIDENT WITH A CHANGE IN THE NEUROFILAMENT FUNCTION.
CC   -!- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z47378; CAA87454.1; -.
DR   InterPro; IPR001664; IF.
DR   Pfam; PF00038; filament; 1.
DR   PROSITE; PS00226; IF; 1.
KW   Intermediate filament; Coiled coil; Neurone; Phosphorylation;
KW   Glycoprotein.
FT   NON_TER       1      1
FT   DOMAIN       <1    197       ROD.
FT   DOMAIN      198    643       TAIL.
FT   DOMAIN       <1     33       COIL 1B.
FT   DOMAIN       34     50       LINKER 12.
FT   DOMAIN       51     72       COIL 2A.
FT   DOMAIN       73     76       LINKER 2.
FT   DOMAIN       77    197       COIL 2B.
FT   CARBOHYD    217    217       O-LINKED (GLCNAC) (BY SIMILARITY).
SQ   SEQUENCE   644 AA;  72450 MW;  030FDAA622889678 CRC64;


  Query Match           11.7%;  Score 149.5;  DB 1;  Length 644;
  Best Local Similarity  31.5%;  Pred. No. 0.012;
  Matches   46;  Conservative   20;  Mismatches   49;  Indels   31;  Gaps    6;

Qy      94 TPNI-------EPQGYSEEEEEEEEEVPAGNAAKAVAKPKAKPAEVKPAVDDEEDESDSD 146
           || |       | :|  ||||||:| | :  | :  ::   |    |   : || |::::
Db     304 TPEIKEEEGEKEEEGQEEEEEEEDEGVKSDQAEEGGSE---KEGSSKNEGEQEEGETEAE 360

Qy     147 G------MDEDDSDGEDSEE----EEPTPKKPASSKKRANETTPKAPVSAKKAKVAVT-- 194
           |       |:    | |||    |||  :      ::|    ||:||   | |   |
Db     361 GEVEEAEAKEEKKTEEKSEEVAAKEEPVTEAKVGKPEKAKSPVPKSPVEEVKPKAEATAG 420

Qy     195 -------PQKTDEKKKGGKAANQSPK 213
                  :| :|:|| ||| :|||
Db     421 KGEQKEEEEKVEEEKK--KAAKESPK 444


RESULT  14
NPM_RAT
ID   NPM_RAT        STANDARD;      PRT;   292 AA.
AC   P13084; Q64269; Q63698;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Nucleophosmin (NPM) (Nucleolar phosphoprotein B23) (Numatrin)
DE   (Nucleolar protein NO38).
GN   NPM1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE (B23.1).
RX   MEDLINE=88330763; PubMed=3417636;
RA   Chang J.-H., Dumbar T.S., Olson M.O.J.;
RT   "cDNA and deduced primary structure of rat protein B23, a nucleolar
RT   protein containing highly conserved sequences.";
RL   J. Biol. Chem. 263:12824-12827(1988).
RN   [2]
RP   SEQUENCE FROM N.A. (B23.2).
RX   MEDLINE=89308567; PubMed=2745414;
RA   Chang J.-H., Olson M.O.J.;
RT   "A single gene codes for two forms of rat nucleolar protein B23
RT   mRNA.";
RL   J. Biol. Chem. 264:11732-11737(1989).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91009311; PubMed=2211699;
RA   Chang J.-H., Olson M.O.J.;
RT   "Structure of the gene for rat nucleolar protein B23.";
RL   J. Biol. Chem. 265:18227-18233(1990).
CC   -!- FUNCTION: ASSOCIATED WITH NUCLEOLAR RIBONUCLEOPROTEIN STRUCTURES
CC       AND BIND SINGLE-STRANDED NUCLEIC ACIDS. IT MAY FUNCTION IN THE
CC       ASSEMBLY AND/OR TRANSPORT OF RIBOSOME.
CC   -!- SUBUNIT: DISULFIDE-LINKED DIMERS UNDER CERTAIN CONDITIONS.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR. GENERALLY NUCLEOLAR, BUT IS
CC       TRANSLOCATED TO THE NUCLEOPLASM IN CASE OF SERUM STARVATION OR
CC       TREATMENT WITH ANTICANCER DRUGS.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; B23.1 (SHOWN HERE) AND B23.2;
CC       ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- PTM: PHOSPHORYLATED.
CC   -!- PTM: THE N-TERMINUS IS BLOCKED.
CC   -!- SIMILARITY: BELONGS TO THE NUCLEOPLASMIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J03969; AAA40796.1; -.
DR   EMBL; M37039; AAA41730.1; -.
DR   EMBL; M37035; AAA41730.1; JOINED.
DR   EMBL; M37036; AAA41730.1; JOINED.
DR   EMBL; M37037; AAA41730.1; JOINED.
DR   EMBL; M37038; AAA41730.1; JOINED.
DR   EMBL; M37041; AAA41731.1; -.
DR   EMBL; M37035; AAA41731.1; JOINED.
DR   EMBL; M37036; AAA41731.1; JOINED.
DR   EMBL; M37037; AAA41731.1; JOINED.
DR   EMBL; M37038; AAA41731.1; JOINED.
DR   EMBL; M37040; AAA41731.1; JOINED.
DR   EMBL; M25062; AAA40795.1; -.
DR   EMBL; J04943; AAA40794.1; -.
DR   EMBL; J04944; AAA40793.1; -.
DR   PIR; A28939; A28939.
DR   InterPro; IPR004301; Nucleoplasmin.
DR   Pfam; PF03066; Nucleoplasmin; 1.
KW   Nuclear protein; Phosphorylation; RNA-binding; Alternative splicing.
FT   DOMAIN        1      9       MET-RICH.
FT   DOMAIN      120    132       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      152    157       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   DOMAIN      158    187       ASP/GLU-RICH (HIGHLY ACIDIC).
FT   DOMAIN      190    196       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   MOD_RES     125    125       PHOSPHORYLATION.
FT   MOD_RES     185    185       PHOSPHORYLATION (BY CK2) (POTENTIAL).
FT   MOD_RES     198    198       PHOSPHORYLATION.
FT   VARSPLIC    256    292       GGSLPKVEAKFINYVKNCFRMTDQEAIQDLWQWRKSL ->
```

```
FT                                AH (IN ISOFORM B23.2).
SQ   SEQUENCE   292 AA;  32560 MW;  1372A474F9ED2457 CRC64;


  Query Match             11.6%;  Score 149;  DB 1;  Length 292;
  Best Local Similarity   27.9%;  Pred. No. 0.0055;
  Matches   39;  Conservative   26;  Mismatches   69;  Indels    6;  Gaps    2;

Qy      83 GKGSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGNAAKAVAKPKAKPAEVKPAVDDEEDE 142
           | | |:  |     :|    ||:|:||: ::   :  ::       | : |  :|:::||
Db     105 GSGPVHISGQHLVAVEEDAESEDEDEEDVKLLGMSGKRSAPGGGNKVPQKKVKLDEDDDE 164

Qy     143 SDSDGMDEDDSDGEDSEEEEPTPKKPASSKKRANET----TPKAPVSAKKAKVAVTPQKT 198
           | |  |::| | :| :|||   | |   ||   :|      |:  : |  | :    |
Db     165 DDEDDEDDEDDDDDDFDEEETEEKVPV--KKSVRDTPAKNAQKSNQNGKDLKPSTPRSKG 222

Qy     199 DEKKKGGKAANQSPKSASQV 218
           |  |  :   ::||  | |
Db     223 QESPKKQEKTPKTPKGPSSV 242


RESULT  15
NP14_RAT
ID   NP14_RAT      STANDARD;      PRT;   704 AA.
AC   P41777;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Nucleolar phosphoprotein p130 (Nucleolar 130 kDa protein) (140 kDa
DE   nucleolar phosphoprotein) (Nopp140) (Nucleolar and coiled-body
DE   phosphoprotein 1).
GN   NOLC1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 292-309 AND 563-601.
RC   TISSUE=Liver;
RX   MEDLINE=92323542; PubMed=1623516;
RA   Meier U.T., Blobel G.;
RT   "Nopp140 shuttles on tracks between nucleolus and cytoplasm.";
RL   Cell 70:127-138(1992).
RN   [2]
RP   INTERACTION WITH NOP5 AND FIBRILLARIN.
RX   MEDLINE=20143579; PubMed=10679015;
RA   Yang Y., Isaac C., Wang C., Dragon F., Pogacic V., Meier U.T.;
RT   "Conserved composition of mammalian box H/ACA and box C/D small
RT   nucleolar ribonucleoprotein particles and their interaction with the
RT   common factor Nopp140.";
RL   Mol. Biol. Cell 11:567-577(2000).
CC   -!- FUNCTION: RELATED TO NUCLEOLOGENESIS, MAY PLAY A ROLE IN THE
CC       MAINTENANCE OF THE FUNDAMENTAL STRUCTURE OF THE FIBRILLAR CENTER
CC       AND DENSE FIBRILLAR COMPONENT IN THE NUCLEOLUS. IT HAS INTRINSIC
CC       GTPASE AND ATPASE ACTIVITIES. MAY PLAY AN IMPORTANT ROLE IN
CC       TRANSCRIPTION CATALYZED BY RNA POLYMERASE I (BY SIMILARITY).
CC   -!- SUBUNIT: Interacts with Dkc1/Nap57, Nop5/Nap65 and fibrillarin.
CC   -!- SUBCELLULAR LOCATION: SHUTTLES ON CURVILINEAR TRACKS BETWEEN
CC       NUCLEOLUS AND CYTOPLASM. THESE TRACKS EXTEND FROM THE DENSE
CC       FIBRILLAR COMPONENT OF THE NUCLEOLUS ACROSS THE NUCLEOPLASM TO
CC       A LIMITED NUMBER OF NUCLEAR PORE COMPLEXES.
CC   -!- PTM: THIS PROTEIN UNDERGOES RAPID AND MASSIVE PHOSPHORYLATION AND
CC       DEPHOSPHORYLATION ON CK-II AND PKC SITES. NOPP140 IS ONE OF THE
CC       MOST PHOSPHORYLATED PROTEINS IN THE CELL.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M94287; AAA41718.1; -.
DR   EMBL; M94288; AAA41719.1; -.
KW   Nuclear protein; Phosphorylation; Repeat; GTP-binding; ATP-binding.
FT   DOMAIN       84    570       11 X 12 AA APPROXIMATE REPEATS OF AN
FT                                ACIDIC SERINE CLUSTER.
FT   REPEAT        84     95       ACIDIC SERINE CLUSTER 1.
FT   REPEAT       127    138       ACIDIC SERINE CLUSTER 2.
FT   REPEAT       170    181       ACIDIC SERINE CLUSTER 3.
FT   REPEAT       231    242       ACIDIC SERINE CLUSTER 4.
FT   REPEAT       274    285       ACIDIC SERINE CLUSTER 5.
FT   REPEAT       335    346       ACIDIC SERINE CLUSTER 6.
FT   REPEAT       373    384       ACIDIC SERINE CLUSTER 7.
FT   REPEAT       434    445       ACIDIC SERINE CLUSTER 8.
FT   REPEAT       479    490       ACIDIC SERINE CLUSTER 9.
FT   REPEAT       524    535       ACIDIC SERINE CLUSTER 10.
FT   REPEAT       559    570       ACIDIC SERINE CLUSTER 11.
FT   MOD_RES      567    567       PHOSPHORYLATION (BY CK2).
FT   VARIANT      150    150       MISSING (IN NOPP140B).
SQ   SEQUENCE   704 AA;  73562 MW;  14DF1BF2DE483EA3 CRC64;


  Query Match             11.4%;  Score 145.5;  DB 1;  Length 704;
  Best Local Similarity   25.0%;  Pred. No. 0.023;
  Matches   52;  Conservative   22;  Mismatches   65;  Indels   69;  Gaps    5;

Qy     103 SEEEEEEEEEVPAGNAAKAVAKPKAKPAEVKPAVDDEEDESDSDGMDEDDSDGEDSEEEE 162
           |:   |||:: |   |   |:|  |||| ||   :   | ||||  :::      |  :
Db     339 SDSSSEEEKKTP---AKTVVSKTPAKPAPVKKKAESSSDSSDSDSSEDEAPAKPVSATKS 395

Qy     163 P----------------TPKKPA--------------------SSKKRANE---TTPK 181
           |                |||:||                    ||:: | :   ||||
Db     396 PLSKPAVTPKPPAAKAVATPKQPAGSGQKPQSRKADSSSSEEESSSSEEEATKKSVTTPK 455

Qy     182 APVSAKKA------------------------KVAVTPQKTDEKKKGGKAANQSPKSAS 216
           | |:|| |                        :   || |   |||   ||   |
Db     456 ARVTAKAAPSLPAKQAPRAGGDSSSDSESSSSEEEKKTPPKPPAKKKAAGAAVPKPTPVK 515

Qy     217 QVSCGSCKKTFNSGNALESHNKAKHAAA 244
           : :  |   : :| :: |   |   : |
Db     516 KAAAESSSSSSSSEDSSEEEKKKPKSKA 543


Search completed: May  2, 2002, 22:31:08
Job time: 1414 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: May 2, 2002, 22:03:59 ; Search time 113.15 Seconds
(without alignments)
374.580 Million cell updates/sec

Title: US-09-645-337-6
Perfect score: 1279
Sequence: 1 MEFWGIEVKSGKPVTVTPEE.....TFNSGNALESHNKAKHAAAK 245

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1279	100.0	245	10	Q9FVE6	Q9fve6 arabidopsis
2	1274	99.6	257	10	O22238	O22238 arabidopsis
3	600.5	47.0	306	10	Q9FNJ6	Q9fnj6 arabidopsis
4	594	46.4	305	10	O49209	O49209 arabidopsis
5	530.5	41.5	303	10	Q9M4U5	Q9m4u5 zea mays (m
6	513.5	40.1	294	10	Q9M4T4	Q9m4t4 arabidopsis
7	511.5	40.0	297	10	Q9M4T5	Q9m4t5 oryza sativ
8	498.5	39.0	296	10	Q9LZR5	Q9lzr5 arabidopsis
9	493	38.5	300	10	Q9M4U4	Q9m4u4 zea mays (m
10	482	37.7	307	10	O24591	O24591 zea mays (m
11	324.5	25.4	286	10	Q94F81	Q94f81 zea mays (m
12	306	23.9	203	10	Q9M4T3	Q9m4t3 arabidopsis
13	179	14.0	477	10	Q93ZG9	Q93zg9 arabidopsis
14	165.5	12.9	990	13	Q91803	Q91803 xenopus lae
15	164	12.8	487	10	Q9STK2	Q9stk2 arabidopsis
16	162.5	12.7	707	11	Q99K50	Q99k50 mus musculu
17	158.5	12.4	399	11	Q99LI4	Q99li4 mus musculu
18	158.5	12.4	513	5	Q95PU8	Q95pu8 chironomus
19	157	12.3	294	4	Q96AT6	Q96at6 homo sapien
20	157	12.3	346	5	Q9NLA3	Q9nla3 asterina pe
21	157	12.3	1110	13	Q91255	Q91255 petromyzon
22	155.5	12.2	444	11	Q9CT46	Q9ct46 mus musculu
23	155	12.1	179	4	Q9NX34	Q9nx34 homo sapien
24	155	12.1	214	4	Q9BTI9	Q9bti9 homo sapien
25	155	12.1	259	4	Q9BYG9	Q9byg9 homo sapien
26	155	12.1	294	4	Q96EA5	Q96ea5 homo sapien
27	153.5	12.0	265	4	Q96DC4	Q96dc4 homo sapien
28	153.5	12.0	635	10	Q40363	Q40363 medicago sa
29	151.5	11.8	715	11	Q9QZX1	Q9qzx1 rattus norv
30	150	11.7	606	4	Q9Y2D9	Q9y2d9 homo sapien
31	148	11.6	296	13	Q91802	Q91802 xenopus lae
32	146.5	11.5	798	5	O45181	O45181 caenorhabdi
33	145.5	11.4	312	5	Q26711	Q26711 trypanosoma
34	144	11.3	257	11	Q9DAY9	Q9day9 mus musculu
35	144	11.3	285	11	Q9DAV3	Q9dav3 mus musculu
36	144	11.3	507	3	Q00880	Q00880 haematonect
37	142.5	11.1	305	3	Q9C2F7	Q9c2f7 neurospora
38	140.5	11.0	649	10	Q9SU25	Q9su25 arabidopsis
39	140.5	11.0	798	13	Q90307	Q90307 carassius a
40	140.5	11.0	1881	16	Q9L7Q2	Q9l7q2 streptococc
41	138.5	10.8	688	5	Q9VNX6	Q9vnx6 drosophila
42	138.5	10.8	971	5	Q9XVS4	Q9xvs4 caenorhabdi
43	138	10.8	418	4	Q96J17	Q96j17 homo sapien
44	138	10.8	700	4	Q9BUV3	Q9buv3 homo sapien
45	138	10.8	705	13	Q06459	Q06459 xenopus lae

ALIGNMENTS

RESULT 1
Q9FVE6
ID Q9FVE6 PRELIMINARY; PRT; 245 AA.
AC Q9FVE6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PUTATIVE HISTONE DEACETYLASE.
GN HD2A.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Wu K., Tian L., Malik K., Brown D., Miki B.;
RT "Functional analysis of HD2 histone deacetylase homologs in
RT Arabidopsis thaliana.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF195545; AAG28472.1; -.
DR InterPro; IPR000822; Znf-C2H2.
DR SMART; SM00355; ZnF_C2H2; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
KW DNA-binding; Metal-binding; Zinc-finger.
SQ SEQUENCE 245 AA; 26372 MW; E8E4F7E0A0D24F28 CRC64;

Query Match 100.0%; Score 1279; DB 10; Length 245;
Best Local Similarity 100.0%; Pred. No. 7.6e-89;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGNAAK 120
```

```
Db      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGNAAK 120

Qy     121 AVAKPKAKPAEVKPAVDDEEDESDSDGMDEDDSDGEDSEEEEPTPKKPASSKKRANETTP 180

Db     121 AVAKPKAKPAEVKPAVDDEEDESDSDGMDEDDSDGEDSEEEEPTPKKPASSKKRANETTP 180

Qy     181 KAPVSAKKAKVAVTPQKTDEKKKGGKAANQSPKSASQVSCGSCKKTFNSGNALESHNKAK 240

Db     181 KAPVSAKKAKVAVTPQKTDEKKKGGKAANQSPKSASQVSCGSCKKTFNSGNALESHNKAK 240

Qy     241 HAAAK 245

Db     241 HAAAK 245


RESULT   2
O22238
ID   O22238      PRELIMINARY;      PRT;   257 AA.
AC   O22238;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE HISTONE DEACETYLASE.
GN   T32N15.8.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   de la Bastide M.R., Parnell L.D., Kaplan N., Gnoj L., Hameed A.,
RA   Schutz K., Hasegawa A., Gottesman T., Shohdy N., Granat S., Jensen K.,
RA   Johnson A.F., Lodhi M., Dedhia N., Martienssen R., McCombie W.R.;
RT   "A. thaliana BAC T32N15 from chromosome V.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC002534; AAB70032.1; -.
DR   InterPro; IPR000822; Znf-C2H2.
DR   Pfam; PF00096; zf-C2H2; 1.
DR   SMART; SM00355; ZnF_C2H2; 1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
KW   DNA-binding; Metal-binding; Zinc-finger.
SQ   SEQUENCE   257 AA;  27698 MW;  C44043876E6889D5 CRC64;


  Query Match           99.6%;  Score 1274;  DB 10;  Length 257;
  Best Local Similarity  100.0%;  Pred. No. 1.9e-88;
  Matches  244;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60

Db       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSEEEEEEEEVPAGNAAK 120

Db      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSEEEEEEEEVPAGNAAK 120

Qy     121 AVAKPKAKPAEVKPAVDDEEDESDSDGMDEDDSDGEDSEEEEPTPKKPASSKKRANETTP 180

Db     121 AVAKPKAKPAEVKPAVDDEEDESDSDGMDEDDSDGEDSEEEEPTPKKPASSKKRANETTP 180

Qy     181 KAPVSAKKAKVAVTPQKTDEKKKGGKAANQSPKSASQVSCGSCKKTFNSGNALESHNKAK 240

Db     181 KAPVSAKKAKVAVTPQKTDEKKKGGKAANQSPKSASQVSCGSCKKTFNSGNALESHNKAK 240

Qy     241 HAAA 244

Db     241 HAAA 244


RESULT   3
Q9FNJ6
ID   Q9FNJ6      PRELIMINARY;      PRT;   306 AA.
AC   Q9FNJ6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HISTONE DEACETYLASE-LIKE PROTEIN.
GN   AT5G22650 OR MDJ22.7.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=COLUMBIA;
RX   MEDLINE=98069011; PubMed=9405937;
RA   Kotani H., Nakamura Y., Sato S., Kaneko T., Asamizu E., Miyajima N.,
RA   Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. II.
RT   Sequence features of the regions of 1,044,062 bp covered by thirteen
RT   physically assigned P1 clones.";
RL   DNA Res. 4:291-300(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
RA   Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA   Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA   Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA   Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA   Ecker J., Theologis A., Davis R.W.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB006699; BAB11671.1; -.
DR   EMBL; AY059893; AAL24375.1; -.
SQ   SEQUENCE   306 AA;  32348 MW;  A19274D43BDD142C CRC64;


  Query Match           47.0%;  Score 600.5;  DB 10;  Length 306;
  Best Local Similarity  50.5%;  Pred. No. 1.1e-37;
  Matches  151;  Conservative   24;  Mismatches   65;  Indels   59;  Gaps   14;

Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60

Db       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASL-DCTVKSGESVVLSVTVGGAKLVIGTLSQ 59

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGY-SEEEEEEEEVPA---- 115

Db      60 DKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPT 119

Qy     116 -----GNAAKAV----AKPKAKPAEVKPAVD----DEEDESDSD----------GM--D 149

Db     120 AVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMDVD 179

Qy     150 EDDSDG---EDS--EEEEPTPKKPAS-SKKRANETTPKAPVSAKKAKVAV----TPQKTD 199

Db     180 EDDSDDDEEEDSEDEEEEEETPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTE 239

Qy     200 EKKKGGKA--------------ANQSPKSASQVSCG-SCKKTFNSGNAL-ESHNKAKH 241

Db     240 EKKKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGGNNNKKPFNSGKQFGGSNNKGSN 298


RESULT   4
O49209
ID   O49209      PRELIMINARY;      PRT;   305 AA.
AC   O49209;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE HISTONE DEACETYLASE.
GN   HD2 OR HD2B.
OS   Arabidopsis thaliana (Mouse-ear cress).
```

```
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Dangl M., Haas H., Loidl P.;
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Wu K., Tian L., Malik K., Brown D., Miki B.;
RT   "Functional analysis of HD2 histone deacetylase homologs in
RT   Arabidopsis thaliana.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF044914; AAC02539.1; -.
DR   EMBL; AF195546; AAG28473.1; -.
SQ   SEQUENCE   305 AA;  32219 MW;  58AA6D5764930944 CRC64;

  Query Match             46.4%;  Score 594;  DB 10;  Length 305;
  Best Local Similarity   50.7%;  Pred. No. 3.3e-37;
  Matches  151;  Conservative   26;  Mismatches   63;  Indels   58;  Gaps   15;

Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
           |||||: |       ||||| |:|:||||| :|  | || | | | ||   ||:||||
Db       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASL-DCTVKSGESVVLSVTVGGAKLVIGTLSQ 59

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGY-SEEEEEEEEEVPA---- 115
             :  ||:  ||||||||||||:  | :|:|:|||:|||| :  | ::|:  | |||
Db      60 DKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPT 119

Qy     116 -----GNAAKAV----AKPKAKPAEVKPAVD----DEEDESDSD-----------GM--D 149
                ||| ||     ||||||||||| :    ||||||| :            ||  |
Db     120 AVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMDVD 179

Qy     150 EDDSDG---EDS--EEEEPTPKKPAS-SKKRANETTPKAPVSAKKAKVAV----TPQKTD 199
           |||||    |||  |||| |||||   :||| ||:  | ||| |||| |     |||||:
Db     180 EDDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTE 239

Qy     200 EK----------KKGGKA---ANQSPKSASQVSCG-SCKKTFNSGNAL-ESHNKAKH 241
           :|          |||||:   |||||| | | | : || |||     |:||  :
Db     240 KKKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGGNNNKKPFNSGKQFGGSNNKGSN 297


RESULT   5
Q9M4U5
ID   Q9M4U5      PRELIMINARY;      PRT;   303 AA.
AC   Q9M4U5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HISTONE DEACETYLASE 2 ISOFORM B.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Dangl M., Brosch G., Haas H., Loidl P., Lusser A.;
RT   "Molecular characterization of type-2 histone deacetylases in higher
RT   plants.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF254072; AAF68624.1; -.
DR   InterPro; IPR000822; Znf-C2H2.
DR   SMART; SM00355; ZnF_C2H2; 1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
KW   DNA-binding; Metal-binding; Zinc-finger.
SQ   SEQUENCE   303 AA;  32613 MW;  708E4627101BB67C CRC64;

  Query Match             41.5%;  Score 530.5;  DB 10;  Length 303;
  Best Local Similarity   42.8%;  Pred. No. 2e-32;
  Matches  130;  Conservative   29;  Mismatches   82;  Indels   63;  Gaps    9;

Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
           |||||:||| |  |   |  | ::|||||:|||  :|| :   ::||| :: | :||||
Db       1 MEFWGLEVKPGSTVKCEPGHGFILHVSQAALGE--SKKSDSALMYVKVDDKKLAIGTLSI 58

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNI---EPQGYSEEEEEEEEEVPA-- 115
             :  ||:  ||||:|||||||||    ||:| |||       :      |:|||| :|
Db      59 DKYPQIQFDLVFNKEFELSHTSKTTSVFFSGYKVEQPIEGDEMDLDSEDEEEELNIPVIK 118

Qy     116 -----------GNAAKAVAKPKAKPA---EVKPAVD-DEEDESDSDGMDEDDSD------ 154
                      |  ||||   :| :   | |   | |: || :||  ||||||
Db     119 ENGKADGKEEQKNQEKAVAATASKSSLGLEKKSKDDSDDSDEDESDDSDEDDSDDSDEGE 178

Qy     155 -----------------GEDSEEEEPTPKKPASSKKRANETTPKAPVSAKKAKVAVTP-Q 196
                           :| ||| |||||| : |||  |   | |:| |||||  | |
Db     179 GLSPDEGDDDSSDEDDTSDDDEEETPTPKKPEAGKKRGAENALKTPLSDKKAKVATPPAQ 238

Qy     197 KTDEKK----------KGGKAAN------QSPKSASQVSCGSCKKTFNSGNALESHNKA 239
           ||  ||           ||   ||      :||||   | | || |||||  ||::|:||
Db     239 KTGGKKGATHVATPHPAKGKTPANNDKLTEKSPKSGGSVPCKSCSKTFNSEMALQAHSKA 298

Qy     240 KHAA 243
           || |
Db     299 KHGA 302


RESULT   6
Q9M4T4
ID   Q9M4T4      PRELIMINARY;      PRT;   294 AA.
AC   Q9M4T4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE HISTONE DEACETYLASE HD2C (AT5G03740/F17C15_160).
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Dangl M., Brosch G., Haas H., Loidl P., Graessle S., Lusser A.;
RT   "Molecular characterization of type-2 histone deacetylases in higher
RT   plants.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=COLUMBIA;
RX   MEDLINE=97471969; PubMed=9330910;
RA   Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
RA   Miyajima N., Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
RT   features of the 1.6 Mb regions covered by twenty physically assigned
RT   P1 clones.";
RL   DNA Res. 4:215-230(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Cheuk R., Chen H., Kim C.J., Meyers M.C., Shinn P., Banh J.,
RA   Bowser L., Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y.,
RA   Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA   Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M.,
RA   Palm C.J., Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M.,
RA   Seki M., Southwick A., Toriumi M., Yamada K., Yu G., Shinozaki K.,
RA   Davis R.W., Theologis A., Ecker J.R.;
RT   "Arabidopsis cDNA clones.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF255712; AAF70197.1; -.
DR   EMBL; AB005235; BAB08599.1; -.
DR   EMBL; AF372089; AAK49605.1; -.
```

```
DR   InterPro; IPR000822; Znf-C2H2.
DR   SMART; SM00355; ZnF_C2H2; 1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
KW   DNA-binding; Metal-binding; Zinc-finger.
SQ   SEQUENCE   294 AA;  31830 MW;  7B69F3BE8324C1D1 CRC64;


  Query Match             40.1%;  Score 513.5;  DB 10;  Length 294;
  Best Local Similarity   40.5%;  Pred. No. 3.6e-31;
  Matches  121;  Conservative   33;  Mismatches   84;  Indels   61;  Gaps    8;

Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
           |||||:|||:|||: : |    |:|:|| :||| ||   | : |:| ||:  |::||||
Db       1 MEFWGVEVKNGKPLHLDPGLDRLVHISQVALGESKNNVTEPIQLYVTVGSDKLLIGTLSH 60

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGNAAK 120
           |  |||  ::| :: | |||||  |||:| |||    :|    | |:  :::: |    |
Db      61 EKFPQLSTEIVLERNFALSHTWKNGSVFFSGYKVDASDP----EPEDLIDDQLEAA-GFK 115

Qy     121 AVAKPKAKPA-------EVKPAVDDEEDESDSDGMDEDDS------------------- 153
           |  |  ||          :||   ||: | |: |  |:|||
Db     116 AAPKSAAKQVNFQLPNEDVKAKQDDDADGSEEDSSDDDDSENSGDEEEEKVTAESDSEED 175

Qy     154 ----DGEDSEEEEPTPKKPASSKKRANE-TTPKAPVSAKKAKVAVTPQKTDEKKK----- 203
               | ||   || |||||   |||: |  : | | | ||||  |||||||  ||
Db     176 DSSDDEEDDSSEEETPKKPEEPKKRSAEPNSSKNPASNKKAKF-VTPQKTDSKKPHVHVA 234

Qy     204 -------------GGKAA-----NQSPKSASQVSCGSCKKTFNSGNALESHNKAKHAAA 244
                       ||        |:||||    | || :|| |   |:|| ||||:||
Db     235 TPHPSKQAGKNSGGGSTGETSKQQQTPKSAGAFGCKSCTRTFTSEMGLQSHTKAKHSAA 293


RESULT   7
Q9M4T5
ID   Q9M4T5      PRELIMINARY;      PRT;   297 AA.
AC   Q9M4T5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE HISTONE DEACETYLASE HD2.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Dangl M., Brosch G., Haas H., Loidl P., Graessle S., Lusser A.;
RT   "Molecular characterization of type-2 histone deacetylases in higher
RT   plants.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF255711; AAF70196.1; -.
DR   InterPro; IPR000822; Znf-C2H2.
DR   SMART; SM00355; ZnF_C2H2; 1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
KW   DNA-binding; Metal-binding; Zinc-finger.
SQ   SEQUENCE   297 AA;  32502 MW;  EEB3AAC0D2C41C63 CRC64;


  Query Match             40.0%;  Score 511.5;  DB 10;  Length 297;
  Best Local Similarity   42.3%;  Pred. No. 5.1e-31;
  Matches  126;  Conservative   36;  Mismatches   79;  Indels   57;  Gaps   10;

Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
           |||||:||| |: |   ||:   :|:||||:||| | |  :   ::||  :| ||:||||
Db       1 MEFWGLEVKPGQTVKCEPEDERFLHLSQAALGESK-KGSDNAVMYVKTDDQKLVIGTLSA 59

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKT--PNIEPQ-----GYSEEEEEEEEEV 113
           :  ||:  |||||||||||||    ||:| |||   |  | :          |:|||||: :
Db      60 DKFPQIQFDLVFDKEFELSHTSKTASVFFSGYKVSQPAEEDEMDFDSEEVEDEEEEEKII 119

Qy     114 PA----------GNAAKAVAKPKAKPAEVKPAVDDEEDESDS---------------DG 147
           ||          |  |   |  :  :: | ||:|:::|: ||               |
Db     120 PAPRANGKVEGKENEQKKQGKTDSSASKSKAAVNDDDDDDDSAEDDSQDEDLSPEDDDDD 179

Qy     148 MDEDDS--DGEDSEEEEPTPKKPASSKKRANETTPKAPVSAKKAKVAVTP--QKTDEKK- 202
             ||||  | ||  :|| ||||| : |::  |   | | | ||||:| ||  ||| :||
Db     180 SSEDDSSEDDEDESDEEDTPKKPETGKRKVAEIVLKTPSSDKKAKIA-TPSGQKTGDKKG 238

Qy     203 ----------------KGGKAANQSPKS-ASQVSCGSCKKTFNSGNALESHNKAKHAA 243
                           |:  :|||    :|| || ||||  ||:||:|||| |
Db     239 VHVATPHPAKQASKTPVNDKSKEKSPKSGGGSISCKSCSKTFNSEMALQSHSKAKHPA 296


RESULT   8
Q9LZR5
ID   Q9LZR5      PRELIMINARY;      PRT;   296 AA.
AC   Q9LZR5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HISTONE DEACETYLASE-LIKE PROTEIN.
GN   F17C15_160.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W., Rudd S.,
RA   Lemcke K., Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL162506; CAB82939.1; -.
DR   InterPro; IPR000822; Znf-C2H2.
DR   SMART; SM00355; ZnF_C2H2; 1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
KW   DNA-binding; Metal-binding; Zinc-finger.
SQ   SEQUENCE   296 AA;  32129 MW;  4B2DAABEAA1F1E9D CRC64;


  Query Match             39.0%;  Score 498.5;  DB 10;  Length 296;
  Best Local Similarity   39.5%;  Pred. No. 4.9e-30;
  Matches  121;  Conservative   31;  Mismatches   81;  Indels   73;  Gaps    9;

Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
           |||||:|||:|||: : |    |:|:|| :||| ||   | : |:| ||:  |::||||
Db       1 MEFWGVEVKNGKPLHLDPGLDRLVHISQVALGESKNNVTEPIQLYVTVGSDKLLIGTLSH 60

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGNAAK 120
           |  |||  ::| :: | |||||  |||:| |||    :|    ::: |      ||   |
Db      61 EKFPQLSTEIVLERNFALSHTWKNGSVFFSGYK-----PEDLIDDQLE-----AAGFKGK 110

Qy     121 -------AVAKPKAKPA-------EVKPAVDDEEDESDSDGMDEDDS------------- 153
                  |  |  ||          :||   ||: | |: |  |:|||
Db     111 WGLLYPPAAPKSAAKQVNFQLPNEDVKAKQDDDADGSEEDSSDDDDSENSGDEEEEKVTA 170

Qy     154 -----------DGEDSEEEEPTPKKPASSKKRANE-TTPKAPVSAKKAKVAVTPQKTDEK 201
                      | ||   || |||||   |||: |  : | | | ||||  |||||||  |
Db     171 ESDSEEDDSSDDEEDDSSEEETPKKPEEPKKRSAEPNSSKNPASNKKAKF-VTPQKTDSK 229

Qy     202 KK-----------------GGKAA-----NQSPKSASQVSCGSCKKTFNSGNALESHNK 238
           |                  ||        |:||||    | || :|| |   |:|| |
Db     230 KPHVHVATPHPSKQAGKNSGGGSTGETSKQQQTPKSAGAFGCKSCTRTFTSEMGLQSHTK 289

Qy     239 AKHAAA 244
           |||:||
Db     290 AKHSAA 295
```

```
RESULT   9
Q9M4U4
ID   Q9M4U4      PRELIMINARY;      PRT;   300 AA.
AC   Q9M4U4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HISTONE DEACETYLASE 2 ISOFORM C.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Dangl M., Brosch G., Haas H., Loidl P., Lusser A.;
RT   "Molecular characterization of type-2 histone deacetylases in higher
RT   plants.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF254073; AAF68625.1; -.
DR   InterPro; IPR000822; Znf-C2H2.
DR   SMART; SM00355; ZnF_C2H2; 1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
KW   DNA-binding; Zinc-finger.
SQ   SEQUENCE   300 AA;  32457 MW;  62F6C4148EB1E1FE CRC64;


  Query Match             38.5%;  Score 493;  DB 10;  Length 300;
  Best Local Similarity   42.1%;  Pred. No. 1.3e-29;
  Matches  128;  Conservative   35;  Mismatches   71;  Indels   70;  Gaps   13;

Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
           |||||:||| |  |    |  | ::|:||||:|||  :|| :   ::||: :| | :||||
Db       1 MEFWGLEVKPGSTVKCEPGYGFVLHLSQAALGE--SKKSDNALMYVKIDDQKLAIGTLSV 58

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSE-------EEEEEEEEV 113
           :  | :  ||:||||||||||    ||:| |||      | : |       |:|:||  |
Db      59 DKNPHIQFDLIFDKEFELSHTSKTTSVFFTGYKV----EQPFEEDEMDLDSEDEDEELNV 114

Qy     114 PA----GNA--------AKAVAKP-KAKPAEVKPAVDDEEDESDSDGMDEDDSD------ 154
           ||    | |          |||| | |: |   |   ||: || ::|  |||::|
Db     115 PAVKENGKADEKKQKSQEKAVAAPSKSSPDSKKSKDDDDSDEDETDDSDEDETDDSDEGL 174

Qy     155 --------------GEDSEEEEPTPKKPASSKKRANETTP-KAPVSAKKAKVAVTP--Q 196
                         :| ||: |||||   |||| |::  | |:| |||||| ||  |
Db     175 SPEEGDDDSSDEDDTSDDEEEDTPTPKKPEVGKKRAAESSVLKTPLSDKKAKVA-TPSSQ 233

Qy     197 KTDEKK-----------KGGKAAN-----QSPKSA--SQVSCGSCKKTFNSGNALESHN 237
           ||  ||            ||    |     :|||||  | | | || |:| |  | ::|:
Db     234 KTGGKKGAAVHVATPHPAKGKTIVNNDKSVKSPKSAPKSGVPCKSCSKSFISETAPQAHS 293

Qy     238 KAKH 241
           ||||
Db     294 KAKH 297


RESULT  10
O24591
ID   O24591      PRELIMINARY;      PRT;   307 AA.
AC   O24591;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NUCLEOLAR HISTONE DEACETYLASE HD2-P39.
GN   HD2.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CUZCO 251;
RX   MEDLINE=97349336; PubMed=9204905;
RA   Lusser A., Brosch G., Loidl A., Haas H., Loidl P.;
RT   "Identification of maize histone deacetylase HD2 as an acidic
RT   nucleolar phosphoprotein.";
RL   Science 277:88-91(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CUZCO;
RA   Brosch G., Lusser A., Loidl P.;
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U82815; AAB63262.1; -.
DR   EMBL; AF026917; AAC61674.1; -.
DR   InterPro; IPR000822; Znf-C2H2.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
KW   DNA-binding; Zinc-finger.
SQ   SEQUENCE   307 AA;  33238 MW;  9CFF6B036D274FC1 CRC64;


  Query Match             37.7%;  Score 482;  DB 10;  Length 307;
  Best Local Similarity   39.0%;  Pred. No. 8.8e-29;
  Matches  121;  Conservative   38;  Mismatches   79;  Indels   72;  Gaps   11;

Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
           |||||:||| |  |    |  | ::|:|||:|||  :|| :   ::||: :| | :||||
Db       1 MEFWGLEVKPGSTVKCEPGYGFVLHLSQAALGE--SKKSDNALMYVKIDDQKLAIGTLSV 58

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSE-------EEEEEEEEV 113
           :  | :  ||:||||||||||    ||:| |||      | : |       |:|:||  |
Db      59 DKNPHIQFDLIFDKEFELSHTSKTTSVFFTGYKV----EQPFEEDEMDLDSEDEDEELNV 114

Qy     114 PA----GNA--------AKAVAKP-KAKPAEVKPAVDDEEDESDSDGMDEDDSD------ 154
           |    | |          |||| | |: |   |   ||: || ::|  |||::|
Db     115 PVVKENGKADEKKQKSQEKAVAAPSKSSPDSKKSKDDDDSDEDETDDSDEDETDDSDEGL 174

Qy     155 --------------GEDSEEEEPTPKKPASSKKR-ANETTPKAPVSAKKAKVAVTP--Q 196
                         :| ||: |||||   ||| |  :  | |:| |||||| ||  |
Db     175 SSEEGDDDSSDEDDTSDDEEEDTPTPKKPEVGKKRPAESSVLKTPLSDKKAKVA-TPSSQ 233

Qy     197 KTDEKK--------------------KGGKAANQSPKSASQVSCGSCKKTFNSGNALES 235
           ||  ||                    | |:   :|||   | |  | |:| | ||::
Db     234 KTGGKKGAAVHVATPHPAKGKTIVNNDKSVKSPKSAPKSGGSVPCKPCSKSFISETALQA 293

Qy     236 HNKAKHAAAK 245
           |::||  |::
Db     294 HSRAKMGASE 303


RESULT  11
Q94F81
ID   Q94F81      PRELIMINARY;      PRT;   286 AA.
AC   Q94F81;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HD2 TYPE HISTONE DEACETYLASE HDA106.
GN   HDA106.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. B73;
RA   Chandler V.L., Kaeppler S.M., Kaeppler H.F., Cone K.C.;
RT   "Sequences from the Plant Chromatin Consortium (NSF Plant Genome
RT   project 9975930).";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. B73;
```

```
RA   Bergstrom D., Springer N.M., Schmitt L.T., Guthrie E., Sidorenko L.,
RA   Selinger D., Kaeppler S.M., Cone K.C.;
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF384033; AAK67143.1; -.
SQ   SEQUENCE   286 AA;  30808 MW;  E295E5CDFE4DEF1A CRC64;


  Query Match             25.4%;  Score 324.5;  DB 10;  Length 286;
  Best Local Similarity   35.7%;  Pred. No. 5.8e-17;
  Matches  102;  Conservative   38;  Mismatches   87;  Indels   59;  Gaps   11;

Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60
           ||  | ||| |  |:    :|::||:|||:||| | |  |   | | : :: |||||||
Db       1 MEVGGQEVKPGATVSCKVGDGLVIHLSQAALGESK-KASENAILSVNIDDKKLVLGTLSV 59

Qy      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSEEEEEEEE----EVPAG 116
           |  ||: |||||||:||| |     ||:| |||:|    :  | |:  :||    ::|
Db      60 EKHPQISCDLVFDKDFELPHNSKTRSVFFRGYKSPVPLFESNSGEDSSDEELKTDQIPLQ 119

Qy     117 N------AAKAVAKPK--------AKPAEVKPAVDDEED----ESDSDGMDEDDSDGED- 157
           |      |||  ||          |   :  : ||::|    :||::  :||||  ||
Db     120 NNEIKISAAKVPAKDDDDDVFIILAMMMMIYSSDDDDDDFTTSDSDNEMSEEDDSSDEDE 179

Qy     158 -SEEEEPT-------------------------PKKP--ASSKKRANETTPKAPVSAKK 188
            |||:: :                         |||      |||| :   :||   ||
Db     180 MSEEDDSSDEDEMSGGADPSDDSSDESGSEHTSAPKKTDVVVGKKRAIKA--EAPY-GKK 236

Qy     189 AKVAVTPQKTDEKKKGGKAANQS---PKSASQVSCGSCKKTFNSGN 231
           ||   : ||| :|      | ||   |   |: :  : ||:  ||:
Db     237 AKSEQSSQKTGDKASTSHPAKQSIKTPADKSRKTPTADKKSPKSGS 282


RESULT  12
Q9M4T3
ID   Q9M4T3      PRELIMINARY;      PRT;   203 AA.
AC   Q9M4T3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   PUTATIVE HISTONE DEACETYLASE HD2D.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Dangl M., Brosch G., Haas H., Loidl P., Graessle S., Lusser A.;
RT   "Molecular characterization of type-2 histone deacetylases in higher
RT   plants.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF255713; AAF70198.1; -.
SQ   SEQUENCE   203 AA;  22651 MW;  69A1DA5941D7A293 CRC64;


  Query Match             23.9%;  Score 306;  DB 10;  Length 203;
  Best Local Similarity   36.8%;  Pred. No. 9.9e-16;
  Matches   86;  Conservative   28;  Mismatches   80;  Indels   40;  Gaps   12;

Qy       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGEC-KNKKGEFVPLHVKVGN--QNLVLGT 57
           |||||||:| |||  |  ::| ::| || :||:  | || |   ::|||:|:     ::|
Db       1 MEFWGIEIKPGKPFKVIQKDGFMVHASQVTLGDVEKVKKDETFAVYVKIGDDENGFMIGN 60

Qy      58 LSTENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSEEEEEEEEEVPAGN 117
           || :  ||   ||    |||:||     ||| :||:|   :      :|| : : |:
Db      61 LS-QKFPQFSIDLYLGHEFEISHN-STSSVYLIGYRT--FDAFDELDEEIDSDSELDEYM 116

Qy     118 AAKAVAKPKAKPAEVKPAVDDEEDESDSD--GMDEDD--SDGEDSEEEEPTPKKPASSKK 173
             :  | |:    |: |   :|:||||||  |:||||  || || | | |    | | |
Db     117 EQQIAALPQ---NEINP---EEDDESDSDEMGLDEDDDSSDEEDVEAEAPLKVAPPSKK- 169

Qy     174 RANETTPKAPVSAKKAKVAVTPQKTDEKKKGGKAANQS-PKSASQVSCG-SCKK 225
                               |    | ||||    | |      || ||||
Db     170 -------------------MPNGAFEIAKGGKKNKSSGGKKRCPFPCGPSCKK 203


RESULT  13
Q93ZG9
ID   Q93ZG9      PRELIMINARY;      PRT;   477 AA.
AC   Q93ZG9;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   AT4G25340/T30C3_20.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA   Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA   Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA   Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA   Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA   Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA   Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA   Ecker J.R.;
RT   "Arabidopsis cDNA clones.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY057543; AAL09783.1; -.
SQ   SEQUENCE   477 AA;  52223 MW;  10C405953B17675E CRC64;


  Query Match             14.0%;  Score 179;  DB 10;  Length 477;
  Best Local Similarity   27.4%;  Pred. No. 8.8e-06;
  Matches   83;  Conservative   33;  Mismatches   87;  Indels  100;  Gaps   15;

Qy       1 MEFWGIEVKSGKPVTVTP--EEGILIHVSQASLGECKNKKGEFVPLHVKVGNQ-NLVLGT 57
           | |||:||| |||    |  |:|  |||:||:||   :|:   :     :|::  : | :
Db       1 MGFWGLEVKPGKPQAYNPKNEQG-KIHVTQATLGTGLSKEKSVI--QCSIGDKAPIALCS 57

Qy      58 LSTENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSEEEEEEEEVPAGN 117
           |    |     :| || : |        | |      :|   |: |  :::|:
Db      58 LLPNKIECCPLNLEFDDDDE--------PVEFTVTGDRSIHLSGFLEYYQDDED------ 103

Qy     118 AAKAVAKPKAKPAEVKPAVDDEEDESDSDGMD-----EDDSDGEDSEEEE---------- 162
                               | | || ||||:|     ||||   ||||:|
Db     104 ------------------DYEHDEDDSDGIDVGESEEDDSCEYDSEEDEQLDEFEDFLD 144

Qy     163 ---------PTPK------------KPASSKKRANETTPKAPVS----AKKAKVAV---- 193
                   ||             |||   | | :|  |:  |    |||  ||:
Db     145 SNLERYRNAAAPKSGVIIEEIEDEEKPAKDNK-AKQTKKKSQASEGENAKKQIVAIEGAH 203

Qy     194 --TPQKTDEKKKG-----GKAANQSPKSASQV-----SCGSCK----KTFNSGNALESHN 237
             :  || : |      ||::     |  ::       || |     |      :  || |
Db     204 VPVLESEDEDEDGLPIPKGKSSEVENASGEKMVVDNDEQGSNKKRKAKAAEQDDGQESAN 263

Qy     238 KAK 240
           |:|
Db     264 KSK 266


RESULT  14
Q91803
ID   Q91803      PRELIMINARY;      PRT;   990 AA.
AC   Q91803;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NUCLEOLAR PHOSPHOPROTEIN.
GN   XNOPP180.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
```

```
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96019267; PubMed=7593294;
RA   Cairns C., McStay B.;
RT   "Identification and cDNA cloning of a Xenopus nucleolar
RT   phosphoprotein, xNopp180, that is the homolog of the rat nucleolar
RT   protein Nopp140.";
RL   J. Cell Sci. 108:3339-3347(1995).
DR   EMBL; X88927; CAA61368.1; -.
SQ   SEQUENCE   990 AA;  102025 MW;  E62054C86A3B4713 CRC64;


  Query Match             12.9%;  Score 165.5;  DB 13;  Length 990;
  Best Local Similarity   30.7%;  Pred. No. 0.0002;
  Matches   55;  Conservative   17;  Mismatches   64;  Indels   43;  Gaps    7;

Qy     103 SEEEEEEEEEVPAGN--AAKAVAKPKAKPAEVKPAVDDEEDESDSDGMDEDDSD------ 154
           |  ||| : : | |   |||| | ||  |  |           ||| | ||
Db     703 SSSEEEGKSKQPTGKSPAAKATAPPKKNPVAVNKDKPSSSSSSDSSGDDEKQKPKQAAAA 762

Qy     155 ---GEDSEEEEPTPKKPASSK------------KRANE-------TTPKAPVSAKKAKV 191
               : ::  :||||| |||              |: |        |||||  :|||
Db     763 KDVKQGAKAAKPTPKKAASSSSEDSSSDEDVSKAKKTNTAVSKSPVTTPKAVPAAKKESS 822

Qy     192 AVTPQKTDEKKKG-----GKAANQSPKSASQVSCGSCKKTFNSGNALESHNKAKHAAAK 245
           : :    |||: |      | || :||:|:       | :  :|    |   ||   : |
Db     823 SESSDSEDEKQGGKNTSTTKIANSTPKAAA----AECSEESSSS---EDEGKANGTSGK 874


RESULT  15
Q9STK2
ID   Q9STK2      PRELIMINARY;      PRT;   487 AA.
AC   Q9STK2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 53.3 KDA PROTEIN.
GN   T30C3.20 OR AT4G25340.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bevan M., Rose M., Hempel S., Entian K.-D., Bancroft I., Mewes H.W.,
RA   Mayer K.F.X., Lemcke K., Schueller C.;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Rose M., Hempel S., Entian K.-D., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL079350; CAB45512.1; -.
DR   EMBL; AL161563; CAB81345.1; -.
DR   HSSP; Q00688; 1PBK.
DR   InterPro; IPR001179; FKBP_PPIase.
DR   Pfam; PF00254; FKBP; 1.
DR   PROSITE; PS00454; FKBP_PPIASE_2; 1.
DR   PROSITE; PS50059; FKBP_PPIASE_3; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   487 AA;  53290 MW;  5B2FA21570BC0AC6 CRC64;


  Query Match             12.8%;  Score 164;  DB 10;  Length 487;
  Best Local Similarity   26.5%;  Pred. No. 0.00012;
  Matches   83;  Conservative   33;  Mismatches   87;  Indels  110;  Gaps   16;

Qy       1 MEFW----------GIEVKSGKPVTVTP--EEGILIHVSQASLGECKNKKGEFVPLHVKV 48
           | ||          |:||| |||    |  |:|  |||:||:||   :|:   :     :
Db       1 MGFWGSDFVLLTYAGLEVKPGKPQAYNPKNEQG-KIHVTQATLGTGLSKEKSVI--QCSI 57

Qy      49 GNQ-NLVLGTLSTENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGYSEEEE 107
           |::  : | :|    |     :| || : |         | |     :|   |: | :
Db      58 GDKAPIALCSLLPNKIECCPLNLEFDDDDE--------PVEFTVTGDRSIHLSGFLEYYQ 109

Qy     108 EEEEEVPAGNAAKAVAKPKAKPAEVKPAVDDEEDESDSDGMD-----EDDSDGEDSEEEE 162
           ::|:                          | | || ||||:|     ||||   ||||:|
Db     110 DDED-----------------------DYEHDEDDSDGIDVGESEEDDSCEYDSEEDE 144

Qy     163 -------------------PTPK------------KPASSKKRANETTPKAPVS----AK 187
                               ||             |||   | | :|  |:  |    ||
Db     145 QLDEFEDFLDSNLERYRNAAAPKSGVIIEEIEDEEKPAKDNK-AKQTKKKSQASEGENAK 203

Qy     188 KAKVAV------TPQKTDEKKKG-----GKAANQSPKSASQV-----SCGSCK----KTF 227
           |  ||:       :  || : |     ||::     |  ::        || |     |
Db     204 KQIVAIEGAHVPVLESEDEDEDGLPIPKGKSSEVENASGEKMVVDNDEQGSNKKRKAKAA 263

Qy     228 NSGNALESHNKAK 240
             :  || ||:|
Db     264 EQDDGQESANKSK 276


Search completed: May  2, 2002, 22:30:22
Job time: 1583 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 20:13:40 ; Search time 5336.92 Seconds
(without alignments)
4775.886 Million cell updates/sec

Title: US-09-645-337-7
Perfect score: 1218
Sequence: 1 gtctttcgcttctaaaaaaa.....aaaaaaaaaagggcggccgc 1218

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1218	100.0	1218	6	AX139020	AX139020 Sequence
2	1116.4	91.7	1187	8	AF195546	AF195546 Arabidops
3	1099.4	90.3	1172	8	AY059893	AY059893 Arabidops
4	847.4	69.6	918	8	AF044914	AF044914 Arabidops
5	309.4	25.4	77363	8	AB006699	AB006699 Arabidops
6	159.2	13.1	977	8	AY065396	AY065396 Arabidops
7	156.2	12.8	914	8	AF195545	AF195545 Arabidops
8	156.2	12.8	939	6	AX139018	AX139018 Sequence
9	146.4	12.0	894	8	AF255711	AF255711 Oryza sat
10	112.8	9.3	903	8	AF254073	AF254073 Zea mays
11	112	9.2	1307	6	AR168371	AR168371 Sequence
12	111.6	9.2	1191	6	AR168369	AR168369 Sequence
13	111.2	9.1	1245	6	AR168370	AR168370 Sequence
14	111.2	9.1	1283	6	AR168368	AR168368 Sequence
15	110	9.0	1169	8	ZMU82815	U82815 Zea mays nu
16	107.4	8.8	912	8	AF254072	AF254072 Zea mays
17	100.6	8.3	1124	8	AF384033	AF384033 Zea mays
c 18	95	7.8	101371	8	AC002534	AC002534 Arabidops
19	90.4	7.4	1236	8	AF372889	AF372889 Arabidops
c 20	85.4	7.0	7218	6	I66494	I66494 Sequence 14
21	84.4	6.9	885	8	AF255712	AF255712 Arabidops
22	68.6	5.6	2486	8	AF026917	AF026917 Zea mays
23	68.4	5.6	1397	10	BC003489	BC003489 Mus muscu
24	68.4	5.6	1408	10	BC005628	BC005628 Mus muscu
25	66.8	5.5	1127	10	AB025582	AB025582 Mus muscu
c 26	65	5.3	209213	2	AL611930	AL611930 Mus muscu
c 27	63.8	5.2	194731	2	AC079365	AC079365 Mus muscu
c 28	62	5.1	132323	8	AP003416	AP003416 Oryza sat
29	62	5.1	174218	2	AC091071	AC091071 Oryza sat
c 30	61.6	5.1	170676	2	AC023448	AC023448 Homo sapi
31	60.8	5.0	75475	8	AB005235	AB005235 Arabidops
32	60.8	5.0	102897	8	ATF17C15	AL162506 Arabidops
33	60.8	5.0	201237	2	AL606932	AL606932 Mus muscu
c 34	60.6	5.0	55296	2	AC101779	AC101779 Mus muscu
35	60.4	5.0	70266	2	AC102233	AC102233 Mus muscu
36	60.4	5.0	205402	2	AL645584	AL645584 Mus muscu
c 37	60	4.9	181238	2	AC094176	AC094176 Rattus no
c 38	60	4.9	203942	2	AL671331	AL671331 Mus muscu
39	59.8	4.9	73132	9	HS833B2	AL023800 Homo sapi
40	59.6	4.9	178820	9	AC020552	AC020552 Homo sapi
41	59.6	4.9	181230	2	AL358876	AL358876 Homo sapi
42	59.4	4.9	204783	2	AL645857	AL645857 Mus muscu
43	59.2	4.9	195028	30	AC024739	AC024739 Homo sapi
44	59	4.8	139665	2	AP000865	AP000865 Homo sapi
45	59	4.8	205638	2	AC040927	AC040927 Mus muscu

ALIGNMENTS

RESULT 1
AX139020
LOCUS AX139020 1218 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 7 from Patent EP1094112.
ACCESSION AX139020
VERSION AX139020.1 GI:14274701
KEYWORDS .
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1218)
AUTHORS Wu,K., Miki,B.L., Tian,L. and Brown,D.C.
TITLE Repressing gene expression in plants
JOURNAL Patent: EP 1094112-A 7 25-APR-2001;
The Minister of Agriculture and Agri-Food (CA)
FEATURES Location/Qualifiers
source 1. .1218
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
BASE COUNT 395 a 237 c 296 g 290 t
ORIGIN

```
  Query Match             100.0%;  Score 1218;  DB 6;  Length 1218;
  Best Local Similarity   100.0%;  Pred. No. 8.6e-276;
  Matches 1218;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 gtctttcgcttctaaaaaaaaacctaacaacctctcttctctcttcctcgttcaacaaca 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GTCTTTCGCTTCTAAAAAAAAACCTAACAACCTCTCTTCTCTCTTCCTCGTTCAACAACA 60

Qy      61 atggagttctggggagttgcggtgacaccaaaaaacgctactaaggtgactcctgaagaa 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ATGGAGTTCTGGGGAGTTGCGGTGACACCAAAAAACGCTACTAAGGTGACTCCTGAAGAA 120

Qy     121 gacagccttgtccacatttctcaggcttcacttgactgcacagtgaaatctggagaatct 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GACAGCCTTGTCCACATTTCTCAGGCTTCACTTGACTGCACAGTGAAATCTGGAGAATCT 180

Qy     181 gtggttttgagtgtgactgttggtggggctaaacttgttattggaacactttcacaagac 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GTGGTTTTGAGTGTGACTGTTGGTGGGGCTAAACTTGTTATTGGAACACTTTCACAAGAC 240

Qy     241 aagttccctcagattagctttgatttggtttttgataaagagtttgagctttcacacagc 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AAGTTCCCTCAGATTAGCTTTGATTTGGTTTTTGATAAAGAGTTTGAGCTTTCACACAGC 300

Qy     301 ggtaccaaagcaaatgttcatttcattggctacaaatcccccaacatcgagcaggatgac 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GGTACCAAAGCAAATGTTCATTTCATTGGCTACAAATCCCCCAACATCGAGCAGGATGAC 360

Qy     361 ttcactagttcggatgatgaggatgttcctgaagctgttcctgctcctgcccctactgct 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTCACTAGTTCGGATGATGAGGATGTTCCTGAAGCTGTTCCTGCTCCTGCCCCTACTGCT 420

Qy     421 gttactgccaacggaaatgctggagcagctgttgtcaaggctgacacaaagccaaaggcc 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTTACTGCCAACGGAAATGCTGGAGCAGCTGTTGTCAAGGCTGACACAAAGCCAAAGGCC 480

Qy     481 aaacctgccgaagtgaagcctgcagaagagaagcctgaatcagacgaggaagatgagtct 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 AAACCTGCCGAAGTGAAGCCTGCAGAAGAGAAGCCTGAATCAGACGAGGAAGATGAGTCT 540

Qy     541 gatgatgaagatgaagtctgaagaggatgatgactcttgagaaaggaatggatggttgat 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GATGATGAAGATGAAGTCTGAAGAGGATGATGACTCTTGAGAAAGGAATGGATGGTTGAT 600

Qy     601 gaagatgactcagatgatgacgaggaggaaggattctgaggatgaagaagaggaggagct 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GAAGATGACTCAGATGATGACGAGGAGGAAGGATTCTGAGGATGAAGAAGAGGAGGAGCT 660

Qy     661 cttaagaagcctgagccaatcaacaagaagaggccaaatgaatctgtatccaaaacaccc 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 CTTAAGAAGCCTGAGCCAATCAACAAGAAGAGGCCAAATGAATCTGTATCCAAAACACCC 720

Qy     721 gtctctggaaagaaggcaaaaccagcagcagcaccagcttctactcctcagaagacagaa 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GTCTCTGGAAAGAAGGCAAAACCAGCAGCAGCACCAGCTTCTACTCCTCAGAAGACAGAA 780

Qy     781 gagaagaagaaaggaggacacaccgccacaccacacccagctaagaagggtggaaagtct 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GAGAAGAAGAAAGGAGGACACACCGCCACACCACACCCAGCTAAGAAGGGTGGAAAGTCT 840

Qy     841 cctgtgaatgctaaccagagccccaagtctggaggtcaatcatccggtggtaacaacaac 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CCTGTGAATGCTAACCAGAGCCCCAAGTCTGGAGGTCAATCATCCGGTGGTAACAACAAC 900

Qy     901 aagaagccattcaactcaggcaaacaatttggtggttccaacaacaagggttctaacaag 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 AAGAAGCCATTCAACTCAGGCAAACAATTTGGTGGTTCCAACAACAAGGGTTCTAACAAG 960

Qy     961 ggcaagggaaagggtagagcttaaggacgtggatcaaggagaggttttgggttttcgagt 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GGCAAGGGAAAGGGTAGAGCTTAAGGACGTGGATCAAGGAGAGGTTTTGGGTTTTCGAGT 1020

Qy    1021 agatgatgaaaacacttggaagtgtggttttggatttttatcttatttttattagtataac 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 AGATGATGAAAACACTTGGAAGTGTGGTTTTGGATTTTTATCTTATTTTTATTAGTATAAC 1080

Qy    1081 ttgttatcggatgagctattttgagtatttgcaatttctactttcctatgtaattcagta 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 TTGTTATCGGATGAGCTATTTTGAGTATTTGCAATTTCTACTTTCCTATGTAATTCAGTA 1140

Qy    1141 tatgaatatttgctgaaatgagaaagaagactcgaattgcaaacaaaaaaaaaaaaaaaa 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TATGAATATTTGCTGAAATGAGAAAGAAGACTCGAATTGCAAACAAAAAAAAAAAAAAA 1200

Qy    1201 aaaaaaaagggcggccgc 1218
           ||||||||||||||||||
Db    1201 AAAAAAAAGGGCGGCCGC 1218


RESULT   2
AF195546
LOCUS       AF195546               1187 bp    mRNA    linear   PLN 01-NOV-2000
DEFINITION  Arabidopsis thaliana putative histone deacetylase (HD2B) mRNA,
            complete cds.
ACCESSION   AF195546
VERSION     AF195546.1  GI:11066136
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1187)
  AUTHORS   Wu,K., Tian,L., Malik,K., Brown,D. and Miki,B.
  TITLE     Functional analysis of HD2 histone deacetylase homologs in
            Arabidopsis thaliana
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1187)
  AUTHORS   Wu,K., Tian,L., Malik,K., Brown,D. and Miki,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-OCT-1999) Eastern Cereal and Oilseed Research Centre,
            Agriculture and Agri-Food Canada, Ottawa, Ontario K1A 0C6, Canada
FEATURES             Location/Qualifiers
     source          1. .1187
                     /organism="Arabidopsis thaliana"
                     /cultivar="Columbia"
                     /db_xref="taxon:3702"
                     /clone="133F15"
     gene            1. .1187
                     /gene="HD2B"
     CDS             61. .978
                     /gene="HD2B"
                     /codon_start=1
                     /product="putative histone deacetylase"
                     /protein_id="AAG28473.1"
                     /db_xref="GI:11066137"
                     /translation="MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLS
                     VTVGGAKLVIGTLSQDKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFT
                     SSDDEDVPEAVPAPAPTAVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDES
                     DDEDESEEDDDSEKGMDVDEDDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKT
                     PVSGKKAKPAAAPASTPQKTEKKKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGGN
                     NNKKPFNSGKQFGGSNNKGSNKGKGKGRA"
BASE COUNT      377 a    234 c    288 g    288 t
ORIGIN


  Query Match             91.7%;  Score 1116.4;  DB 8;  Length 1187;
  Best Local Similarity   99.2%;  Pred. No. 6.4e-252;
  Matches 1185;  Conservative    0;  Mismatches    1;  Indels    8;  Gaps    6;

Qy       1 gtctttcgcttctaaaaaaaaacctaacaacctctcttctctcttcctcgttcaacaaca 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GTCTTTCGCTTCTAAAAAAAAACCTAACAACCTCTCTTCTCTCTTCCTCGTTCAACAACA 60
```

```
Qy      61 atggagttctggggagttgcggtgacaccaaaaaaacgctactaaggtgactcctgaagaa 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ATGGAGTTCTGGGGAGTTGCGGTGACACCAAAAAACGCTACTAAGGTGACTCCTGAAGAA 120

Qy     121 gacagccttgtccacatttctcaggcttcacttgactgcacagtgaaatctggagaatct 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GACAGCCTTGTCCACATTTCTCAGGCTTCACTTGACTGCACAGTGAAATCTGGAGAATCT 180

Qy     181 gtggttttgagtgtgactgttggtggggctaaacttgttattggaacactttcacaagac 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GTGGTTTTGAGTGTGACTGTTGGTGGGGCTAAACTTGTTATTGGAACACTTTCACAAGAC 240

Qy     241 aagttccctcagattagctttgatttggtttttgataaagagtttgagctttcacacagc 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AAGTTCCCTCAGATTAGCTTTGATTTGGTTTTTGATAAAGAGTTTGAGCTTTCACACAGC 300

Qy     301 ggtaccaaagcaaatgttcatttcattggctacaaatcccccaacatcgagcaggatgac 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GGTACCAAAGCAAATGTTCATTTCATTGGCTACAAATCCCCCAACATCGAGCAGGATGAC 360

Qy     361 ttcactagttcggatgatgaggatgttcctgaagctgttcctgctcctgcccctactgct 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTCACTAGTTCGGATGATGAGGATGTTCCTGAAGCTGTTCCTGCTCCTGCCCCTACTGCT 420

Qy     421 gttactgccaacggaaatgctggagcagctgttgtcaaggctgacacaaagccaaaggcc 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTTACTGCCAACGGAAATGCTGGAGCAGCTGTTGTCAAGGCTGACACAAAGCCAAAGGCC 480

Qy     481 aaacctgccgaagtgaagcctgcagaagagaagcctgaatcagacgaggaagatgagtct 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 AAACCTGCCGAAGTGAAGCCTGCAGAAGAGAAGCCTGAATCAGACGAGGAAGATGAGTCT 540

Qy     541 gatgatgaagatgaagtctgaagaggatgatgactcttgagaaaggaatggatggttgat 600
           |||||||||||||| |||||||||||||||||||| |||||||||||||||| ||||||
Db     541 GATGATGAAGATG-AGTCTGAAGAGGATGATGACTC-TGAGAAAGGAATGGAT-GTTGAT 597

Qy     601 gaagatgactcagatgatgacgaggaggaaggattctgaggatgaagaagaggaggag-c 659
           |||||||||||||||||||||||||| |||||||||||||||||||||||||||||| |
Db     598 GAAGATGACTCAGATGATGACGAGGAGG-AGGATTCTGAGGATGAAGAAGAGGAGGAGAC 656

Qy     660 tcttaagaagcctgagccaatcaacaagaagaggccaaatgaatctgtatccaaaacacc 719
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     657 TCCTAAGAAGCCTGAGCCAATCAACAAGAAGAGGCCAAATGAATCTGTATCCAAAACACC 716

Qy     720 cgtctctggaaagaaggcaaaaccagcagcagcaccagcttctactcctcagaagacaga 779
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     717 CGTCTCTGGAAAGAAGGCAAAACCAGCAGCAGCACCAGCTTCTACTCCTCAGAAGAC--- 773

Qy     780 agagaagaagaaaggaggacacaccgccacaccacacccagctaagaagggtggaaagtc 839
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     774 AGAGAAGAAGAAAGGAGGACACACCGCCACACCACACCCAGCTAAGAAGGGTGGAAAGTC 833

Qy     840 tcctgtgaatgctaaccagagccccaagtctggaggtcaatcatccggtggtaacaacaa 899
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     834 TCCTGTGAATGCTAACCAGAGCCCCAAGTCTGGAGGTCAATCATCCGGTGGTAACAACAA 893

Qy     900 caagaagccattcaactcaggcaaacaatttggtggttccaacaacaagggttctaacaa 959
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     894 CAAGAAGCCATTCAACTCAGCCAAACAATTTGGTGGTTCCAACAACAAGGGTTCTAACAA 953

Qy     960 gggcaagggaaagggtagagcttaaggacgtggatcaaggagaggttttgggttttcgag 1019
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     954 GGGCAAGGGAAAGGGTAGAGCTTAAGGACGTGGATCAAGGAGAGGTTTTGGGTTTTCGAG 1013

Qy    1020 tagatgatgaaaacacttggaagtgtggttttggatttttatcttatttttattagtataa 1079
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1014 TAGATGATGAAAACACTTGGAAGTGTGGTTTTGGATTTTTATCTTATTTTTATTAGTATAA 1073

Qy    1080 cttgttatcggatgagctattttgagtatttgcaatttctactttcctatgtaattcagt 1139
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1074 CTTGTTATCGGATGAGCTATTTTGAGTATTTGCAATTTCTACTTTCCTATGTAATTCAGT 1133

Qy    1140 atatgaatatttgctgaaatgagaaagaagactcgaattgcaaacaaaaaaaaa 1193
```

```
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1134 ATATGAATATTTGCTGAAATGAGAAAGAAGACTCGAATTGCAAACAAAAAAAAA 1187


RESULT   3
AY059893
LOCUS       AY059893                1172 bp    mRNA    linear   PLN 04-NOV-2001
DEFINITION  Arabidopsis thaliana histone deacetylase-like protein (At5g22650;
            MDJ22.7) mRNA, complete cds.
ACCESSION   AY059893
VERSION     AY059893.1  GI:16649046
KEYWORDS    FLI_CDNA.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1172)
  AUTHORS   Nguyen,M., Karlin-Neumann,G., Southwick,A., Lam,B., Miranda,M.,
            Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,
            Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
            Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
            Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,
            Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-OCT-2001) DNA Sequencing and Technology Center,
            Stanford University, 855 California Avenue, Palo Alto, CA 94304,
            USA
COMMENT     e-mail for correspondence: arab@sequence.stanford.edu

            RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis  Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members carried out the
            sequencing and annotation of the RAFL cDNAs: Nguyen,M.,
            Southwick,A., Karlin-Neumann,G.,  Lam,B., Miranda,M., Palm,C.J.,
            Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K.,
            Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P.,
            Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

            Nguyen,M. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed
            equally to this work.  Shinozaki,K. (RIKEN GSC) and Davis,R.W.
            (SSP/Stanford)  contributed equally to this work as PIs.
FEATURES             Location/Qualifiers
     source          1. .1172
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /chromosome="V"
                     /clone="RAFL08-18-E14"
                     /note="This clone is in pBluescript"
     gene            1. .1172
                     /gene="At5g22650; MDJ22.7"
     CDS             75. .995
                     /gene="At5g22650; MDJ22.7"
                     /codon_start=1
                     /product="histone deacetylase-like protein"
                     /protein_id="AAL24375.1"
                     /db_xref="GI:16649047"
                     /translation="MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLS
                     VTVGGAKLVIGTLSQDKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFT
                     SSDDEDVPEAVPAPAPTAVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDES
                     DDEDESEEDDDSEKGMDVDEDDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKT
                     PVSGKKAKPAAAPASTPQKTEEKKKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGG
                     NNNKKPFNSGKQFGGSNNKGSNKGKGKGRA"
BASE COUNT      362 a    234 c    285 g    291 t
ORIGIN


  Query Match           90.3%;  Score 1099.4;  DB 8;  Length 1172;
  Best Local Similarity 99.5%;  Pred. No. 6.3e-248;
```

```
  Matches 1155;  Conservative    0;  Mismatches    1;  Indels    5;  Gaps    5;

Qy       2 tctttcgcttctaaaaaaaaacctaacaacctctcttctctcttcctcgttcaacaacaa 61
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      16 TCTTTCGCTTCTAAAAAAAAACCTAACAACCTCTCTTCTCTCTTCCTCGTTCAACAACAA 75

Qy      62 tggagttctggggagttgcggtgacaccaaaaaacgctactaaggtgactcctgaagaag 121
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      76 TGGAGTTCTGGGGAGTTGCGGTGACACCAAAAAACGCTACTAAGGTGACTCCTGAAGAAG 135

Qy     122 acagccttgtccacatttctcaggcttcacttgactgcacagtgaaatctggagaatctg 181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     136 ACAGCCTTGTCCACATTTCTCAGGCTTCACTTGACTGCACAGTGAAATCTGGAGAATCTG 195

Qy     182 tggttttgagtgtgactgttggtggggctaaacttgttattggaacactttcacaagaca 241
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     196 TGGTTTTGAGTGTGACTGTTGGTGGGGCTAAACTTGTTATTGGAACACTTTCACAAGACA 255

Qy     242 agttccctcagattagctttgatttggtttttgataaagagtttgagctttcacacagcg 301
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     256 AGTTCCCTCAGATTAGCTTTGATTTGGTTTTTGATAAAGAGTTTGAGCTTTCACACAGCG 315

Qy     302 gtaccaaagcaaatgttcatttcattggctacaaatccccccaacatcgagcaggatgact 361
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     316 GTACCAAAGCAAATGTTCATTTCATTGGCTACAAATCCCCCAACATCGAGCAGGATGACT 375

Qy     362 tcactagttcggatgatgaggatgttcctgaagctgttcctgctcctgccctactgctg 421
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     376 TCACTAGTTCGGATGATGAGGATGTTCCTGAAGCTGTTCCTGCTCCTGCCCCTACTGCTG 435

Qy     422 ttactgccaacggaaatgctggagcagctgttgtcaaggctgacacaaagccaaaggcca 481
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     436 TTACTGCCAACGGAAATGCTGGAGCAGCTGTTGTCAAGGCTGACACAAAGCCAAAGGCCA 495

Qy     482 aacctgccgaagtgaagcctgcagaagagaagcctgaatcagacgaggaagatgagtctg 541
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     496 AACCTGCCGAAGTGAAGCCTGCAGAAGAGAAGCCTGAATCAGACGAGGAAGATGAGTCTG 555

Qy     542 atgatgaagatgaagtctgaagaggatgatgactcttgagaaaggaatggatggttgatg 601
           |||||||||||| |||||||||||||||||||| |||||||||||||||| |||||||
Db     556 ATGATGAAGATG-AGTCTGAAGAGGATGATGACTC-TGAGAAAGGAATGGAT-GTTGATG 612

Qy     602 aagatgactcagatgatgacgaggaggaaggattctgaggatgaagaagaggaggag-ct 660
           |||||||||||||||||||||||||| |||||||||||||||||||||||||||||| ||
Db     613 AAGATGACTCAGATGATGACGAGGAGG-AGGATTCTGAGGATGAAGAAGAGGAGGAGACT 671

Qy     661 cttaagaagcctgagccaatcaacaagaagaggccaaatgaatctgtatccaaaacaccc 720
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     672 CCTAAGAAGCCTGAGCCAATCAACAAGAAGAGGCCAAATGAATCTGTATCCAAAACACCC 731

Qy     721 gtctctggaaagaaggcaaaaccagcagcagcaccagcttctactcctcagaagacagaa 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     732 GTCTCTGGAAAGAAGGCAAAACCAGCAGCAGCACCAGCTTCTACTCCTCAGAAGACAGAA 791

Qy     781 gagaagaagaaagaaggagacacaccgccacaccacacccagctaagaaggqtggaaagtct 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     792 GAGAAGAAGAAAGGAGGACACACCGCCACACCACACCCAGCTAAGAAGGGTGGAAAGTCT 851

Qy     841 cctgtgaatgctaaccagagccccaagtctggaggtcaatcatccggtggtaacaacaac 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     852 CCTGTGAATGCTAACCAGAGCCCCAAGTCTGGAGGTCAATCATCCGGTGGTAACAACAAC 911

Qy     901 aagaagccattcaactcaggcaaacaatttggtggttccaacaacaagggttctaacaag 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     912 AAGAAGCCATTCAACTCAGGCAAACAATTTGGTGGTTCCAACAACAAGGGTTCTAACAAG 971

Qy     961 gqcaagggaaagggtagagcttaaggacgtggatcaaggagaggttttgggttttcgagt 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     972 GGCAAGGGAAAGGGTAGAGCTTAAGGACGTGGATCAAGGAGAGGTTTTGGGTTTTCGAGT 1031

Qy    1021 agatgatgaaaacacttggaagtgtggttttggatttttatcttattttattagtataac 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1032 AGATGATGAAAACACTTGGAAGTGTGGTTTTGGATTTTTATCTTATTTTATTAGTATAAC 1091

Qy    1081 ttgttatcggatgagctattttgagtatttgcaatttctacttcctatgtaattcagta 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1092 TTGTTATCGGATGAGCTATTTTGAGTATTTGCAATTTCTACTTTCCTATGTAATTCAGTA 1151

Qy    1141 tatgaatatttgctgaaatga 1161
           |||||||||||||||||||||
Db    1152 TATGAATATTTGCTGAAATGA 1172


RESULT   4
AF044914
LOCUS       AF044914                 918 bp    mRNA    linear   PLN 11-FEB-1998
DEFINITION  Arabidopsis thaliana putative histone deacetylase (HD2) mRNA,
            complete cds.
ACCESSION   AF044914
VERSION     AF044914.1  GI:2854069
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 918)
  AUTHORS   Dangl,M., Haas,H. and Loidl,P.
  TITLE     Arabidopsis thaliana complete cDNA-sequence homologous to Zea mays
            HD2
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 918)
  AUTHORS   Dangl,M., Haas,H. and Loidl,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JAN-1998) Microbiology, Medical School, University of
            Innsbruck, Fritz-Pregl-Str. 3, Innsbruck 6020, Austria
FEATURES             Location/Qualifiers
     source          1. .918
                     /organism="Arabidopsis thaliana"
                     /cultivar="Columbia"
                     /db_xref="taxon:3702"
                     /clone="133F15T7"
                     /clone_lib="EST-clone 133F15 (T. Newman, Plant Physiol.
                     106, 1241-1255, 1994)"
     gene            1. .918
                     /gene="HD2"
     CDS             1. .918
                     /gene="HD2"
                     /codon_start=1
                     /product="putative histone deacetylase"
                     /protein_id="AAC02539.1"
                     /db_xref="GI:2854070"
                     /translation="MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLS
                     VTVGGAKLVIGTLSQDKFPQISFDLVFDKEPELSHSGTKANVHFIGYKSPNIEQDDFT
                     SSDDEDVPEAVPAPAPTAVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDES
                     DDEDESEEDDDSEKGMDVDEDDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKT
                     PVSGKKAKPAAAPASTPQKTEKKKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGGN
                     NNKKPFNSGKQFGGSNNKGSNKGKGKGRA"
BASE COUNT      292 a    194 c    239 g    193 t
ORIGIN


  Query Match             69.6%;  Score 847.4;  DB 8;  Length 918;
  Best Local Similarity   99.0%;  Pred. No. 1e-188;
  Matches  916;  Conservative    0;  Mismatches    1;  Indels    8;  Gaps    6;

Qy      61 atggagttctggggagttgcggtgacaccaaaaaacgctactaaggtgactcctgaagaa 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGGAGTTCTGGGGAGTTGCGGTGACACCAAAAAACGCTACTAAGGTGACTCCTGAAGAA 60

Qy     121 gacagccttgtccacatttctcaggcttcacttgactgcacagtgaaatctggagaatct 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GACAGCCTTGTCCACATTTCTCAGGCTTCACTTGACTGCACAGTGAAATCTGGAGAATCT 120

Qy     181 gtggttttgagtgtgactgttggtggggctaaacttgttattggaacactttcacaagac 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    121 GTGGTTTTGAGTGTGACTGTTGGTGGGGCTAAACTTGTTATTGGAACACTTTCACAAGAC 180

Qy    241 aagttccctcagattagctttgatttggtttttgataaagagtttgagctttcacacagc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAGTTCCCTCAGATTAGCTTTGATTTGGTTTTTGATAAAGAGTTTGAGCTTTCACACAGC 240

Qy    301 ggtaccaaagcaaatgttcatttcattggctacaaatcccccaacatcgagcaggatgac 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGTACCAAAGCAAATGTTCATTTCATTGGCTACAAATCCCCCAACATCGAGCAGGATGAC 300

Qy    361 ttcactagttcggatgatgaggatgttcctgaagctgttcctgctcctgcccctactgct 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TTCACTAGTTCGGATGATGAGGATGTTCCTGAAGCTGTTCCTGCTCCTGCCCCTACTGCT 360

Qy    421 gttactgccaacggaaatgctggagcagctgttgtcaaggctgacacaaagccaaaggcc 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GTTACTGCCAACGGAAATGCTGGAGCAGCTGTTGTCAAGGCTGACACAAAGCCAAAGGCC 420

Qy    481 aaacctgccgaagtgaagcctgcagaagagaagcctgaatcagacgaggaagatgagtct 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AAACCTGCCGAAGTGAAGCCTGCAGAAGAGAAGCCTGAATCAGACGAGGAAGATGAGTCT 480

Qy    541 gatgatgaagatgaagtctgaagaggatgatgactcttgagaaaggaatggatggttgat 600
          |||||||||||||| |||||||||||||||||||||| |||||||||||||||| ||||||
Db    481 GATGATGAAGATG-AGTCTGAAGAGGATGATGACTC-TGAGAAAGGAATGGAT-GTTGAT 537

Qy    601 gaagatgactcagatgatgacgaggaggaaggattctgaggatgaagaagaggaggag-c 659
          |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||| |
Db    538 GAAGATGACTCAGATGATGACGAGGAGG-AGGATTCTGAGGATGAAGAAGAGGAGGAGAC 596

Qy    660 tcttaagaagcctgagccaatcaacaagaagaggccaaatgaatctgtatccaaaacacc 719
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    597 TCCTAAGAAGCCTGAGCCAATCAACAAGAAGAGGCCAAATGAATCTGTATCCAAAACACC 656

Qy    720 cgtctctggaaagaaggcaaaaccagcagcagcaccagcttctactcctcagaagacaga 779
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    657 CGTCTCTGGAAAGAAGGCAAAACCAGCAGCAGCACCAGCTTCTACTCCTCAGAAGAC--- 713

Qy    780 agagaagaagaaaggaggacacaccgccacaccacacccagctaagaagggtggaaagtc 839
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    714 AGAGAAGAAGAAAGGAGGACACACCGCCACACCACACCCAGCTAAGAAGGGTGGAAAGTC 773

Qy    840 tcctgtgaatgctaaccagagccccaagtctggaggtcaatcatccggtggtaacaacaa 899
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    774 TCCTGTGAATGCTAACCAGAGCCCCAAGTCTGGAGGTCAATCATCCGGTGGTAACAACAA 833

Qy    900 caagaagccattcaactcaggcaaacaatttggtggttccaacaacaagggttctaacaa 959
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    834 CAAGAAGCCATTCAACTCAGGCAAACAATTTGGTGGTTCCAACAACAAGGGTTCTAACAA 893

Qy    960 gggcaagggaaagggtagagcttaa 984
          |||||||||||||||||||||||||
Db    894 GGGCAAGGGAAAGGGTAGAGCTTAA 918


RESULT   5
AB006699
LOCUS       AB006699               77363 bp    DNA     linear   PLN 27-DEC-2000
DEFINITION  Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:MDJ22.
ACCESSION   AB006699 BA000015
VERSION     AB006699.1  GI:2351064
KEYWORDS    .
SOURCE      Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui P1
            clone:MDJ22.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (sites)
  AUTHORS   Kotani,H., Nakamura,Y., Sato,S., Kaneko,T., Asamizu,E., Miyajima,N.
            and Tabata,S.
  TITLE     Structural analysis of Arabidopsis thaliana chromosome 5. II.
            Sequence features of the regions of 1,044,062 bp covered by
            thirteen physically assigned P1 clones
  JOURNAL   DNA Res. 4 (4), 291-300 (1997)
  MEDLINE   98069011
REFERENCE   2  (bases 1 to 77363)
  AUTHORS   Nakamura,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-AUG-1997) Yasukazu Nakamura, Kazusa DNA Research
            Institute, Department of Plant Gene Research; 1532-3, Yana,
            Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp,
            Tel:81-438-52-3935, Fax:81-438-52-3934)
COMMENT     Address for correspondence: kaos@kazusa.or.jp
            For the latest information on annotation of this clone, please see
            http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=MDJ22
            Genes with similarity to proteins in the databases are described in
            'product' or 'note' qualifiers.  Genes that have no significant
            protein similarity are described as 'unknown protein'.
            The software programs used to predict genes include: Grail
            (Informatics Group, Oak Ridge National Laboratory,
            http://compbio.ornl.gov/Grail-1.3/),
            GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
            NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
            Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
            SplicePredictor (Volker Brendel, Stanford University,
            http://gremlin1.zool.iastate.edu/cgi-bin/sp.cgi).
            Genes encoding tRNAs are predicted by tRNAscan-SE
            (Sean Eddy, Washington University School of Medicine, St. Louis,
            http://genome.wustl.edu/eddy/tRNAscan-SE/).
            This sequence may not be the entire insert of this clone. It may be
            shorter because we remove overlaps between neighboring submissions.
            The 5' clone is MQJ16 and the 3' clone is K5A21.
FEATURES              Location/Qualifiers
     source           1. .77363
                      /organism="Arabidopsis thaliana"
                      /strain="Columbia"
                      /db_xref="taxon:3702"
                      /chromosome="5"
                      /clone="MDJ22"
                      /clone_lib="Mitsui P1"
     CDS              join(1014. .1043,1182. .1414,1564. .2029,2131. .2316)
                      /note="gene_id:MDJ22.1
                      unknown protein"
                      /codon_start=1
                      /evidence=not_experimental
                      /protein_id="BAB11665.1"
                      /db_xref="GI:10178233"
                      /translation="MGPYLGPMRKILLYLPTKKAVSVLSKRWKSLWLSFPGLDFDSEE
                      FPNYNAFVSFMDRFVDFSRERKSCLHKLKLSIQKSWSDQYRITRLESISLPRLKTVHL
                      QLNVYANDAGLESFISSYPAHEDFSIVRSVDIDDNVNVLRVHSQTLTSLSVEFDLGEW
                      GLGLLCFGQRYLGLWIDDPRFKYLNFKHELSDSKTISNLNSLVKVNFVGFYMSHGVGL
                      SKQQMARNVFTGISRVRDLIISADIFIDCYLKEESLPQFCNLSCLKAEVCSSNVKIRE
                      TLQKPFESCPNLKSIFLVMTCTNPRALT"
     CDS              join(6501. .6675,6836. .7458,7550. .7699,7802. .8110)
                      /note="gene_id:MDJ22.2
                      unknown protein"
                      /codon_start=1
                      /evidence=not_experimental
                      /protein_id="BAB11666.1"
                      /db_xref="GI:10178234"
                      /translation="MEARKEKRICLERPAREDLISKLPNALISHISRIFGQRMLSGLV
                      FYPRGGKVSGVWFPDYVARGNLQHLDVELGGCVYFREYWEVMPVSLYICETLLHLRLY
                      RVSMGNFESVSLPHLKTLCLERNIYPNEASVESLISSCPVLEDLTILRINDNVKVLRV
                      HSQSLTSFSVGYHPGDLIRKYNYYSEKVRENSGLVIDAPRLKYLTFNYERSKSKIMRN
                      LSSLVKVNVLSSFDISSVAGCSEQQMAHNFFTGISRVRDLIISRDMMRLIFYYLKGNS
                      LPQFCNLSYLKTNLSETSLAFLDIFTKLLESCPNLKSIALGLTCFLDDMVEMSVWSVP
                      KCLLSSLEFVEIKNEHPPDDGVLKVARYFVENSVNLKKLALHLHFFFLEGNPAVLNDL
                      LALPRRSSMCQIEVFNVQSGRSRGIF"
     CDS              join(11695. .12657,12751. .12900,12999. .13394)
                      /note="gene_id:MDJ22.3
                      unknown protein"
                      /codon_start=1
                      /evidence=not_experimental
                      /protein_id="BAB11667.1"
                      /db_xref="GI:10178235"
```

/translation="MEENIEKRICVEQLAIEDLISKLPEVLLSQILSYLPTKDIVRTS
VLSKRWKSVWLLIPGLDLDSSEFPHYDTFVDFMNEFLFFSREENPCLHKLKLSIQKNE
NDPSCVTLWTDCVARGKLQHLDVEFGGRVMEREFWEMMPLSLYICKTLLHLRLYRVLL
GNFDQSVDSLPSLKSMCLEENVYSNEASLESLISSCRVLEDLTIVKIDDNVRFLRVHS
QSLTSLSVGYKSYYPGEIYYYYDRDRGNSGLVIDAPRLKYLTFNNDQSKSKTISNLGS
LVKVTILGPIKISRVVGCTEQQMAHKFLTGISRVRYLIVSEDMMEVISSYLKEDSLPQ
FGNLSYLKASVWLSSFDFLDILPKLLESCPNLKSIVLETTCIVDRTKATVERRVSSVP
ECLLSSLEFVEIKNRISVDDGALEVARYFVENSVNLQKVVLRLASSFLRRGNQAVLKD
ILELPRRSSMCQIEVFNALNGHALCFRKNKITGRVFLDYFDVAESYYGAIVS"

CDS complement(join(14209. .14547,14624. .14731,14851. .14949,
15047. .15112,15289. .15381,15541. .15659,15755. .15971,
16065. .16170,16400. .16701))
/note="dbj|BAA92923.1
gene_id:MDJ22.4
similar to unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB11668.1"
/db_xref="GI:10178236"
/translation="MISLPLTTPILPSRCLLHKTRRQNSTRRRLLIRSSSSLQDQFTV
ETTKRVVLVRHGQSTWNEEGRIQGSSDFSVLTKKGESQAEISRQMLIDDSFDVCFTSP
LKRSKKTAEIIWGSRESEMIFDYDLREIDLYSFQGLLKKEGKEKFGEAFKQWQEDPAN
FIIDGHYPVRELWSRARSCWPGILAHESKSVLVVAHNAVNQALLATAIGLGTEYFRSL
LQSNCGVSVLDFIPRADGGSPHVCLNRLNQTPNSPLAGGSSGGRKASKQIILVCHGQG
NNEDSAVINQAANNDQAMNMLGVIHSQKTAELLLDLRVSSIVCSPKTASIESSGVISR
VQEAAGCLGVDNVPHYVKTKQMNELDVESVLRKSNKDNDVIASQLDEEAFSALWNRSE
KAWESLLDELSDEKSNPGEIMVVVGPAMTHISLIAQCLNLTKEALGLFHLDAGSISVI
DFPDGPSSKGVIRCTNYTAHLGRWSIPITKPA"

CDS 21123. .22400
/note="gene_id:MDJ22.5"
/codon_start=1
/evidence=not_experimental
/product="chorismate mutase/prephenate dehydratase-like
protein"
/protein_id="BAB11669.1"
/db_xref="GI:10178237"
/translation="MQTISPAFSCDLKSVIQPNLTAKKARYSHVNGKRVSVRCSYRSE
SFSFPNGVGSSRADWQSSCAILASKVVSAENSSSVAVVNGHSNGSVDLSLVPSKSQHN
GKPGLIQPLTITDLSPAPSHGSTLRVAYQGVPGAYSEAAAGKAYPNSEAIPCDQFDVA
FQAVELWIADRAVLPVENSLGGSIHRNYDLLLRHRLHIVGEVQIPVHHCLLALPGVRT
DCITRVISHPQALAQTEGSLNKLTPKAAIEAFHDTAAAAEYIAANNLHDTAAVASARA
AELYGLQILADGIQDDAGNVTRFLMLARDPIIPRTDRPFKTSIVFAAQEHKGTSVLFK
VLSAFAFRNISLTKIESRPHQNCPVRVVGDENVGTSKHFEYTFYVDFEASMAEARAQN
ALAEVQEYTSFLRVLGSYPMDMTPWSTLPSEDV"

CDS join(25901. .26553,26964. .27035,27152. .27241,27328. .27501,
27723. .27813,27972. .28034,28129. .28408,28466. .28569,
28648. .28774,28862. .29628,29712. .30119)
/note="gene_id:MDJ22.6
unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB11670.1"
/db_xref="GI:10178238"
/translation="MANEELTESQQQEDPSQQLPNADEEKGSDSDSNSDSDASSQSSG
DDFYISESENEAEGDNTIFNYVRPSDIPPDPNANPETNIRRFNRVLDGKRVKRMQEEE
EDKYTFYEDLFDFPRDPERWKEQDLREIWADGPLEMTKPGWDPAWADEDDWDVVNDEI
QEGRDPGIQPFYVPYRKPYPAIPDNHYDIENAKGVVEELDRIEEFLQWVSYIFPDGSS
YEGTVWDDLAQGKGVYIAENGLVRYEGEWLQNDMEGHGVIDVDIPDIEPIPGSKLEAK
MRAEGRIIKRDYMTPEDRKWLEMDVEDSVALTDGNFQVPFYENEEWVTQFGEKPEKGR
YRYAGQWKHSRMHGCGVYEVNERILYGRFYFGELLEEEHGCTVDICALHSGLAEVAAA
KARMFVNKPDGSIAPFLYHFFCLVLCIKYKPRHILGICLIPLDQCRTLSSNHHSVSNI
YFAFLSGNVQLMKVRKKTGHFTVRLIREERGPYGDPQHPYFYEEDDVWMAPGFINQFY
EVPEYWETYVGEVDQEREMWLNSFYKAPLRLPMPAELEHWWENVEVTPEFVLLNKEPE
PDPNDPSKLVQKEDPVILHTPTGRIINYVEDEKHGIRLFWQPPLEEGEEVDPSKVEFL
PLGFDEFYGKEVVVKKEHPIKSFVLGIEKSVKPMLDGLEKWTEEKKKAYEERKEMIQQ
ELELVEAEICLEEAIEDMDEELKKKEQEEEKKTEMGLTEEDEDVLVPVYKEEKVVTAK
EKIQENKQEEKYKDDDDEDDDDGDDDDDDDDDDLGPSSFGSADKGRRNSPFSSSSLS
FASCTLFPAVQSRLESSFLAWKQHRAEPSKVNTGIIKGADTASASIHFPPLSSNNARL
KMGKVANRGCVQRSYGSSRSQSQLMSLSRLLSCNASSSSSPPDSSSSEYLKDSGLWET
PVGDMSVVLSLQIQTKCSDLFAETPAVS"

CDS join(30598. .30610,30871. .30941,31032. .31240,31371. .31450,
31536. .31687,31802. .31882,31983. .32091,32233. .32359,
32454. .32532)
/note="gene_id:MDJ22.7"
/codon_start=1
/evidence=not_experimental
/product="histone deacetylase-like protein"
/protein_id="BAB11671.1"
/db_xref="GI:10178239"
/translation="MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLS
VTVGGAKLVIGTLSQDKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFT
SSDDEDVPEAVPAPAPTAVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDES
DDEDESEEDDDSEKGMDVDEDDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKT
PVSGKKAKPAAAPASTPQKTEEKKKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGG
NNNKKPFNSGKQFGGSNNKGSNKGKGKGRA"

CDS complement(join(32806. .33093,33184. .33324,33408. .34331))
/note="gene_id:MDJ22.8
unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB11672.1"
/db_xref="GI:10178240"
/translation="MSNQGAIRRSGEDRISSLPDHLLSQILSNLPTENAVTTSILSTR
WKDLWLSTPVLDIDIDAFDDATTFISFATRFLDSFKDSCLHKLQISFQMEAVDMWTII
PWIEDAVKRRIQHLEVDSRIDHMIDTLPLTVYLSESLVSLRLHLVMLHRFVFVSLPNL
KVMHLEENIYSYAETMEKFISSCPVLEDLTVVRNVDEATEKVLRVSSQSLNSLKLVID
SSKCWYNDDSDDWKVVIDAPQLVYLSLKDDQSVSFVINNLCSSAKADIKVSFNVSDIW
DLEESFERSNVGKFLTGLSSLRDMTISGTTLKIICHYLKHEPMPQFRNMTRLHAKFYV
CDLEMLPCVLESCPNLKSLVLKLKGEMENEEISLSSSVPKCLQSSLENVEIIRPNYGS
GEEMKLSKYFLENSLVLKKFKLCRDCHSEEQESLVVRELMTFQRCSSACEINVVRFQR
"

CDS complement(join(34670. .34948,35032. .35172,35252. .36163))
/note="gene_id:MDJ22.9
unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB11673.1"
/db_xref="GI:10178241"
/translation="MSTKRGVTSQDSISLLPDDLLCRILSNLPTKVAVRTSVLSKRWK
RFSLSVPLLEFNVSEFHGYYEFARVVHGFLDTSRETCIHKLKLAFEKKQHDRSYLTQW
IHNAVKRKVQHLDIGRWSYLGQELIPHSLYTCETLVSLRLHNVSLPDFDHVSLPRLKT

Query Match 25.4%; Score 309.4; DB 8; Length 77363;
Best Local Similarity 81.3%; Pred. No. 3.8e-62;
Matches 414; Conservative 0; Mismatches 1; Indels 94; Gaps 1;

```
Qy    775 acagaagagaagaagaaagaggagacacaccgccacaccacacccagctaagaagggtgga 834
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  32229 ACAGAAGAGAAGAAGAAAGGAGGACACACCGCCACACCACACCCAGCTAAGAAGGGTGGA 32288

Qy    835 aagtctcctgtgaatgctaaccagagccccaagtctggaggtcaatcatccggtggtaac 894
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  32289 AAGTCTCCTGTGAATGCTAACCAGAGCCCCAAGTCTGGAGGTCAATCATCCGGTGGTAAC 32348

Qy    895 aacaacaagaa---------------------------------------------------- 905
          ||||||||||| 
Db  32349 AACAACAAGAAGTGAGTTGTCTCTTTCTTTATCTTTCTCATAGCCCAAAGATATGAGTGA 32408

Qy    906 ----------------------------------------------gccattcaactcagg 920
                                                         ||||||||||||||| 
Db  32409 AAATGTTAATTGGTAATGATGTGTAATGGTATTGTGTGTGTACAGGCCATTCAACTCAGG 32468

Qy    921 caaacaatttggtggttccaacaacaagggttctaacaagggcaagggaaagggtagagc 980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  32469 CAAACAATTTGGTGGTTCCAACAACAAGGGTTCTAACAAGGGCAAGGGAAAGGGTAGAGC 32528

Qy    981 ttaaggacgtggatcaaggagaggttttgggttttcgagtagatgatgaaaacacttggga 1040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  32529 TTAAGGACGTGGATCAAGGAGAGGTTTTGGGTTTTCGAGTAGATGATGAAAACACTTGGA 32588

Qy   1041 agtgtggttttggatttttatcttattttattagtataacttgttatcggatgagctatt 1100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  32589 AGTGTGGTTTTGGATTTTTATCTTATTTTATTAGTATAACTTGTTATCGGATGAGCTATT 32648

Qy   1101 ttgagtatttgcaatttctactttcctatgtaattcagtatatgaatatttgctgaaatg 1160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  32649 TTGAGTATTTGCAATTTCTACTTTCCTATGTAATTCAGTATATGAATATTTGCTGAAATG 32708

Qy   1161 aqaaagaagactcqaattgcaaacaaaaa 1189
          |||||||||||||||||||||||||| |||
Db  32709 AGAAAGAAGACTCGAATTGCAAACACAAA 32737


RESULT  6
AY065396
LOCUS       AY065396                977 bp    mRNA    linear   PLN 11-DEC-2001
DEFINITION  Arabidopsis thaliana putative histone deacetylase (At3g44750) mRNA,
            complete cds.
ACCESSION   AY065396
VERSION     AY065396.1  GI:17529219
KEYWORDS    FLI_CDNA.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 977)
  AUTHORS   Yamada,K., Banh,J., Chang,C.H., Chang,E., Dale,J.M.,
            Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.,
            Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Yu,S., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
            Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
            Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Full Length cDNA of gene At3g44750 (GI:15230496)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 977)
  AUTHORS   Yamada,K., Banh,J., Chang,C.H., Chang,E., Dale,J.M.,
            Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C.,
            Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Yu,S., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
            Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
            Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-DEC-2001) Plant Gene Expression Center, 800 Buchanan
            Street, Albany, CA 94710, USA
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members carried out the
            sequencing and annotation of the RAFL cDNAs: Yamada,K., Banh,J.,
            Chang,C.H., Chang,E., Dale,J.M., Goldsmith,A.D., Lee,J.M.,
            Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C.,
            Yamamura,Y., Yu,G., Yu,S., Bowser,L., Chen,H., Cheuk,R., Jones,T.,
            Karlin-Neumann,G., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M.,
            Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Davis,R.W.,
            Ecker,J.R. and Theologis,A.

            Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to
            this work.  Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC)
            contributed equally to this work as PIs.
FEATURES             Location/Qualifiers
     source          1. .977
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /chromosome="3"
                     /clone="RAFL06-70-K17 (R11479)"
                     /note="This clone is in a modified pBluescript vector
                     (Lambda ZAP) as a XhoI/SstI insert.
                     ecotype: Columbia"
     gene            1. .977
                     /gene="At3g44750"
     5'UTR           1. .62
                     /gene="At3g44750"
     CDS             63. .800
                     /gene="At3g44750"
                     /codon_start=1
                     /evidence-experimental
                     /product="putative histone deacetylase"
                     /protein_id="AAL38837.1"
                     /db_xref="GI:17529220"
                     /translation="MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPL
                     HVKVGNQNLVLGTLSTENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGY
                     SEEEEEEEEEVPAGNAAKAVAKPKAKPAEVKPAVDDEEDESDSDGMDEDDSDGEDSEE
                     EEPTPKKPASSKKRANETTPKAPVSAKKAKVAVTPQKTDEKKKGGKAANQSPKSASQV
                     SCGSCKKTFNSGNALESHNKAKHAAAK"
     3'UTR           801. .977
                     /gene="At3g44750"
BASE COUNT      311 a    183 c    231 g    252 t
ORIGIN


  Query Match           13.1%;  Score 159.2;  DB 8;  Length 977;
  Best Local Similarity  62.4%;  Pred. No. 5.4e-27;
  Matches  267;  Conservative    0;  Mismatches  158;  Indels    3;  Gaps    1;

Qy       2 tctttcgcttctaaaaaaaaacctaacaacctctcttctctcttcctcgttcaacaacaa 61
           | ||  ||| || | |||      ||  |    || ||||  ||| | |    | | |
Db       4 TTTTCAGCTGCTCATAAACCCTAAAAATCCTCTCTTTTTCTCAACCTTGATTCTTAGCCA 63

Qy      62 tggagttctggggagttgcggtgacaccaaaaaacgctactaaggtgactcctgaagaag 121
           ||||||||||||| ||| || | | ||  |||  |   ||  ||||||||||||||||
Db      64 TGGAGTTCTGGGGAATTGAAGTTAAATCAGGAAAGCCAGTTACAGTGACTCCTGAAGAAG 123

Qy     122 acagccttgtccacatttctcaggcttcactt---gactgcacagtgaaatctggagaat 178
            ||  ||| ||||| ||||||||||| || |||   || || | |   ||    ||||| |
Db     124 GCATTCTTATCCACGTTTCTCAGGCATCGCTTGGAGAATGTAAAAACAAGAAGGGAGAGT 183

Qy     179 ctgtggttttgagtgtgactgttggtggggctaaacttgttattggaacactttcacaag 238
           ||||  |||   ||| | |||||       ||  | ||| | ||||| || ||    |
Db     184 TTGTGCCTTTACATGTAAAGGTTGGGAACCAGAACTTGGTTCTGGGAACTCTATCGACTG 243

Qy     239 acaagttccctcagattagctttgatttggtttttgataaagagtttgagctttcacaca 298
           | ||  |||||||| ||  || |||||||| || || || ||||||||||||| ||||
Db     244 AGAACATCCCTCAGCTTTTCTGTGATTTGGTATTCGACAAGGAGTTTGAGCTTTCTCACA 303

Qy     299 gcggtaccaaagcaaatgttcatttcattggctacaaatcccccaacatcgagcaggatg 358
             |    |||| || |||| | ||  |||| |||||| | |||||||| ||||   | |
Db     304 CTTGGGGAAAAGGAAGTGTTTACTTTGTTGGATACAAAACTCCCAACATTGAGCCACAAG 363

Qy     359 acttcactagttcggatgatgaggatgttcctgaagctgttcctgctcctgcccctactg 418
            ||   ||    || || || || |      |||| |  | |||    |||  | | |
Db     364 GCTATTCTGAGGAAGAAGAGGAAGAAGAGGAAGAAGTTCCTGCTGGGAATGCTGCCAAGG 423

Qy     419 ctgttact 426
           |||| ||
Db     424 CTGTAGCT 431


RESULT  7
AF195545
LOCUS       AF195545                914 bp    mRNA    linear   PLN 01-NOV-2000
DEFINITION  Arabidopsis thaliana putative histone deacetylase (HD2A) mRNA,
            complete cds.
ACCESSION   AF195545
VERSION     AF195545.1  GI:11066134
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 914)
  AUTHORS   Wu,K., Tian,L., Malik,K., Brown,D. and Miki,B.
  TITLE     Functional analysis of HD2 histone deacetylase homologs in
            Arabidopsis thaliana
```

```
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 914)
  AUTHORS   Wu,K., Tian,L., Malik,K., Brown,D. and Miki,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-OCT-1999) Eastern Cereal and Oilseed Research Centre,
            Agriculture and Agri-Food Canada, Ottawa, Ontario K1A 0C6, Canada
FEATURES             Location/Qualifiers
     source          1. .914
                     /organism="Arabidopsis thaliana"
                     /cultivar="Columbia"
                     /db_xref="taxon:3702"
                     /clone="211N2"
     gene            1. .914
                     /gene="HD2A"
     CDS             49. .786
                     /gene="HD2A"
                     /codon_start=1
                     /product="putative histone deacetylase"
                     /protein_id="AAG28472.1"
                     /db_xref="GI:11066135"
                     /translation="MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPL
                     HVKVCNQNLVLGTLSTENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGY
                     SEEEEEEEEEVPAGNAAKAVAKPKAKPAEVKPAVDDEEDESDSDGMDEDDSDGEDSEE
                     EEPTPKKPASSKKRANETTPKAPVSAKKAKVAVTPQKTDEKKKGGKAANQSPKSASQV
                     SCGSCKKTFNSGNALESHNKAKHAAAK"
BASE COUNT      281 a    171 c    230 g    232 t
ORIGIN
```

```
  Query Match           12.8%;  Score 156.2;  DB 8;  Length 914;
  Best Local Similarity   64.1%;  Pred. No. 2.7e-26;
  Matches  252;  Conservative    0;  Mismatches  138;  Indels    3;  Gaps    1;

Qy      37 ttctctcttcctcgttcaacaacaatggagttctggggagttgcggtgacaccaaaaaac 96
           || ||||  ||| | |    | | ||||||||||||||||| |||  || | | ||  |||
Db      25 TTTTCTCAACCTTGATTCTTAGCCATGGAGTTCTGGGGAATTGAAGTTAAATCAGGAAAG 84

Qy      97 gctactaaggtgactcctgaagaagacagccttgtccacatttctcaggcttcactt--- 153
             |   ||  |||||||||||||||| ||  ||| ||||| |||||||||| || |||
Db      85 CCAGTTACAGTGACTCCTGAAGAAGGCATTCTTATCCACGTTTCTCAGGCATCGCTTGGA 144

Qy     154 gactgcacagtgaaatctggagaatctgtggttttgagtgtgactgttggtggggctaaa 213
           || || | |   ||    ||||| | ||||  |||   ||| |  |||||        ||
Db     145 GAATGTAAAAACAAGAAGGGAGAGTTTGTGCCTTTACATGTAAAGGTTGGGAACCAGAAC 204

Qy     214 cttgttattggaacactttcacaagacaagttccctcagattagctttgatttggttttt 273
             | ||| | ||||| || ||    || ||  |||||||| ||  || |||||||| ||
Db     205 TTGGTTCTGGGAACTCTATCGACTGAGAACATCCCTCAGCTTTTCTGTGATTTGGTATTC 264

Qy     274 gataaagagtttgagctttcacacagcggtaccaaagcaaatgttcatttcattggctac 333
           || || ||||||||||||||| ||||   |    |||| || |||| | ||  |||| |||
Db     265 GACAAGGAGTTTGAGCTTTCTCACACTTGGGGAAAAGGAAGTGTTTACTTTGTTGGATAC 324

Qy     334 aaatcccccaacatcgagcaggatgacttcactagttcggatgatgaggatgttcctgaa 393
           ||| | |||||||| ||||   | | ||   ||      || || || || |     |||
Db     325 AAAACTCCCAACATTGAGCCACAAGGCTATTCTGAGGAAGAAGAGGAAGAAGAGGAAGAA 384

Qy     394 gctgttcctgctcctgcccctactgctgttact 426
           | |  | |||    |||  | |  |||||  ||
Db     385 GTTCCTGCTGGGAATGCTGCCAAGGCTGTAGCT 417
```

```
RESULT   8
AX139018
LOCUS       AX139018                939 bp    DNA     linear   PAT 30-MAY-2001
DEFINITION  Sequence 5 from Patent EP1094112.
ACCESSION   AX139018
VERSION     AX139018.1  GI:14274700
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 939)
  AUTHORS   Wu,K., Miki,B.L., Tian,L. and Brown,D.C.
  TITLE     Repressing gene expression in plants
  JOURNAL   Patent: EP 1094112-A 5 25-APR-2001;
            The Minister of Agriculture and Agri-Food (CA)
FEATURES             Location/Qualifiers
     source          1. .939
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
BASE COUNT      296 a    175 c    236 g    232 t
ORIGIN
```

```
  Query Match           12.8%;  Score 156.2;  DB 6;  Length 939;
  Best Local Similarity   64.1%;  Pred. No. 2.7e-26;
  Matches  252;  Conservative    0;  Mismatches  138;  Indels    3;  Gaps    1;

Qy      37 ttctctcttcctcgttcaacaacaatggagttctggggagttgcggtgacaccaaaaaac 96
           || ||||  ||| | |    | | ||||||||||||||||| |||  || | | ||  |||
Db      25 TTTTCTCAACCTTGATTCTTAGCCATGGAGTTCTGGGGAATTGAAGTTAAATCAGGAAAG 84

Qy      97 gctactaaggtgactcctgaagaagacagccttgtccacatttctcaggcttcactt--- 153
             |   ||  |||||||||||||||| ||  ||| ||||| |||||||||| || |||
Db      85 CCAGTTACAGTGACTCCTGAAGAAGGCATTCTTATCCACGTTTCTCAGGCATCGCTTGGA 144

Qy     154 gactgcacagtgaaatctggagaatctgtggttttgagtgtgactgttggtggggctaaa 213
           || || | |   ||    ||||| | ||||  |||   ||| |  |||||        ||
Db     145 GAATGTAAAAACAAGAAGGGAGAGTTTGTGCCTTTACATGTAAAGGTTGGGAACCAGAAC 204

Qy     214 cttgttattggaacactttcacaagacaagttccctcagattagctttgatttggttttt 273
             | ||| | ||||| || ||    || ||  |||||||| ||  || |||||||| ||
Db     205 TTGGTTCTGGGAACTCTATCGACTGAGAACATCCCTCAGCTTTTCTGTGATTTGGTATTC 264

Qy     274 gataaagagtttgagctttcacacagcggtaccaaagcaaatgttcatttcattggctac 333
           || || ||||||||||||||| ||||   |    |||| || |||| | ||  |||| |||
Db     265 GACAAGGAGTTTGAGCTTTCTCACACTTGGGGAAAAGGAAGTGTTTACTTTGTTGGATAC 324

Qy     334 aaatcccccaacatcgagcaggatgacttcactagttcggatgatgaggatgttcctgaa 393
           ||| | |||||||| ||||   | | ||   ||      || || || || |     |||
Db     325 AAAACTCCCAACATTGAGCCACAAGGCTATTCTGAGGAAGAAGAGGAAGAAGAGGAAGAA 384

Qy     394 gctgttcctgctcctgcccctactgctgttact 426
           | |  | |||    |||  | |  |||||  ||
Db     385 GTTCCTGCTGGGAATGCTGCCAAGGCTGTAGCT 417
```

```
RESULT   9
AF255711
LOCUS       AF255711                894 bp    mRNA    linear   PLN 15-MAY-2000
DEFINITION  Oryza sativa putative histone deacetylase HD2 mRNA, complete cds.
ACCESSION   AF255711
VERSION     AF255711.1  GI:7804959
KEYWORDS    .
SOURCE      Oryza sativa.
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 894)
  AUTHORS   Dangl,M., Brosch,G., Haas,H., Loidl,P., Graessle,S. and Lusser,A.
  TITLE     Molecular characterization of type-2 histone deacetylases in higher
            plants
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 894)
  AUTHORS   Dangl,M., Haas,H., Loidl,P. and Lusser,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-APR-2000) Microbiology, University of Innsbruck,
            Fritz-Pregl-Str. 3, Innsbruck 6020, Austria
FEATURES             Location/Qualifiers
     source          1. .894
                     /organism="Oryza sativa"
                     /db_xref="dbEST:D15380"
```

```
                     /db_xref="taxon:4530"
     CDS             1. .894
                     /codon_start=1
                     /product="putative histone deacetylase HD2"
                     /protein_id="AAF70196.1"
                     /db_xref="GI:7804960"
                     /translation="MEFWGLEVKPGQTVKCEPEDERFLHLSQAALGESKKGSDNAVMY
                     VKTDDQKLVIGTLSADKFPQIQFDLVFDKEFELSHTSKTASVFFSGYKVSQPAEEDEM
                     DFDSEEVEDEEEEEKIIPAPRANGKVEGKENEQKKQGKTDSSASKSKAAVNDDDDDDD
                     SAEDDSQDEDLSPEDDDDDSSEDDSSEDDEDESDEEDTPKKPETGKRKVAEIVLKTPS
                     SDKKAKIATPSGQKTGDKKGVHVATPHPAKQASKTPVNDKSKEKSPKSGGGSISCKSC
                     SKTFNSEMALQSHSKAKHPAK"
BASE COUNT      289 a    172 c    234 g    199 t
ORIGIN


  Query Match             12.0%;  Score 146.4;  DB 8;  Length 894;
  Best Local Similarity   53.4%;  Pred. No. 5.5e-24;
  Matches  434;  Conservative    0;  Mismatches  351;  Indels   27;  Gaps    5;

Qy      61 atggagttctggggagttgcggtgacaccaaaaaacgctactaaggtgactcctgaagaa 120
           ||||||||||||||  |||  || |  ||   | |  ||   ||       ||||||||
Db       1 ATGGAGTTCTGGGGTCTTGAAGTCAAGCCTGGACAGACTGTCAAATGTGAGCCTGAAGAT 60

Qy     121 gacagccttgtccacatttctcaggcttcacttgactgcacagtgaaatctggagaatct 180
           ||  || || | ||| |||||||||| | ||||      ||  ||||      ||   |
Db      61 GAACGCTTTTTGCACCTTTCTCAGGCTGCTCTTGGGGAATCAAAGAAAGGATCTGACAAT 120

Qy     181 gtggttttgagtgtgactgttggtggggctaaacttgttattggaacactttcacaagac 240
           |  ||  ||  ||| |    || ||     || || || |||||||| || |||   |||
Db     121 GCAGTAATGTATGTTAAAACTGATGATCAAAAGCTAGTCATTGGAACCCTCTCAGCTGAC 180

Qy     241 aagttccctcagattagctttgatttggtttttgataaagagtttgagctttcacacagc 300
           |||||||||| ||    |||||||||| ||||| |||||||||||||| |||||||
Db     181 AAGTTCCCTCAAATCCAGTTTGATTTGGTCTTTGACAAAGAGTTTGAGCTGTCACACACT 240

Qy     301 ggtaccaaagcaaatgttcatttcattggctacaaa---tcccccaacatcgagcaggat 357
           |  |  || | |||    ||   |||||||||   ||||        ||| | |||
Db     241 TCAAAGACTGCTAGTGTGTTCTTTTCTGGCTACAAAGTTTCCCAGCCGGCTGAGGAAGAT 300

Qy     358 gacttcactagttcggatgatgaggatgttcctgaagctgttcctgctcctgccctact 417
           ||  |   |  |     ||| || | ||    ||||| |     |       |  | |
Db     301 GAAATGGATTTTCATTCTGAAGAAGTTGAAGATGAAGAGGAGGAAGAAAAGATCATTCCA 360

Qy     418 gctgttactgccaacggaaatgctggagcagctgttgtcaaggctgacacaaagccaaag 477
           |||   |  || || || | | || ||     |      ||    | | | |    |||
Db     361 GCTCCCAGGGCAAATGGCAAAGTTGAAGGGAAGGAAAATGAGCAGAAAAAACAAGGCAAG 420

Qy     478 gccaaacctgccgaagtgaagcctgcagaagagaagcctgaatcagacgaggaagatgag 537
           |  |  || | |     ||  |    | |     | |||    || || || ||||
Db     421 ACAGATTCTTCAGCTTCAAAATCAAAGGCTGCAGTGAATGACGATGATGATGATGATGAC 480

Qy     538 tctgatgatgaagatgaagtctgaagaggatgatgactcttgagaaaggaatggatggtt 597
           ||  || || |||    ||| | || || |||   |||   ||        |   | |
Db     481 AGTGCAGAGGATGAT----TCTCAGGACGAAGATCTTTCTCCTGA--------GGATGAT 528

Qy     598 gatgaagatgactcagatgatgacgaggaggaaggattctgaggatgaagaagaggagga 657
           |||||  |||   || || |||||      ||||      |     ||     |  | |
Db     529 GATGATGATTCTTCTGAGGATGATTCCAGCGAAGATGATGAGGATGAGAGTGACGAGGAA 588

Qy     658 gctcttaagaagcctgagccaatcaacaagaagaggccaaatgaatctgtatccaaaaca 717
           |  |  |  |||          |     |||| ||    |  |||| ||| |  || |||
Db     589 GATACTCCCAAGAAGCCAGAGACTGGAAAGAGGAAAGTAGCTGAAATTGTGTTGAAGACA 648

Qy     718 cccgtctctggaaagaaggcaaaaccagcagcagcaccagcttctactcctcagaagaca 777
           ||    ||||  ||||| |||||    ||  || |  |||           |||||||||
Db     649 CCTTCGTCTGATAAGAAAGCAAACATTGCTACACCGTCAG---------GCCAGAAGACA 699

Qy     778 gaagagaagaagaaaggaggacacaccgccacaccacacccagctaagaagggtggaaag 837
           |  || ||||||   || |  ||    || || ||||| || || ||| ||| | | |||
Db     700 GGTGACAAGAAG---GGTGTCCATGTAGCAACTCCACATCCGGCAAAGCAGGCTAGCAAG 756

Qy     838 tctcctgtgaatgctaaccagagccccaagtc 869
```

```
           | || |||||||  ||    |    |||||
Db     757 ACCCCCGTGAATGACAAGTCAAAGGAGAAGTC 788


RESULT  10
AF254073
LOCUS       AF254073                903 bp    mRNA    linear   PLN 08-MAY-2000
DEFINITION  Zea mays histone deacetylase 2 isoform c mRNA, complete cds.
ACCESSION   AF254073
VERSION     AF254073.1  GI:7716949
KEYWORDS    .
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 903)
  AUTHORS   Dangl,M., Brosch,G., Haas,H., Loidl,P. and Lusser,A.
  TITLE     Molecular characterization of type-2 histone deacetylases in higher
            plants
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 903)
  AUTHORS   Lusser,A. and Loidl,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-APR-2000) Microbiology, Medical School, University of
            Innsbruck, Fritz-Pregl Str. 3/II, Innsbruck 6020, Austria
FEATURES             Location/Qualifiers
     source          1. .903
                     /organism="Zea mays"
                     /db_xref="taxon:4577"
     CDS             1. .903
                     /note="HD2c"
                     /codon_start=1
                     /product="histone deacetylase 2 isoform c"
                     /protein_id="AAF68625.1"
                     /db_xref="GI:7716950"
                     /translation="MEFWGLEVKPGSTVKCEPGYGFVLHLSQAALGESKKSDNALMYV
                     KIDDQKLAIGTLSVDKNPHIQFDLIFDKEFELSHTSKTTSVFFTGYKVEQPFEEDEMD
                     LDSEDEDEELNVPAVKENGKADEKKQKSQEKAVAAPSKSSPDSKKSKDDDDSDEDETD
                     DSDEDETDDSDEGLSPEEGDDDSSDEDDTSDDEEEDTPTPKKPEVGKKRAAESSVLKT
                     PLSDKKAKVATPSSQKTGGKKGAAVHVATPHPAKGKTIVNNDKSVKSPKSAPKSGVPC
                     KSCSKSFISETAPQAHSKAKHGGK"
BASE COUNT      280 a    188 c    238 g    197 t
ORIGIN


  Query Match             9.3%;  Score 112.8;  DB 8;  Length 903;
  Best Local Similarity   50.8%;  Pred. No. 4.3e-16;
  Matches  360;  Conservative    0;  Mismatches  327;  Indels   21;  Gaps    3;

Qy      61 atggagttctggggagttgcggtgacaccaaaaaacgctactaaggtgactcctgaagaa 120
           ||||||||||||||  | | ||| |  ||      | ||  ||||      |||| | |
Db       1 ATGGAGTTCTGGGGTCTCGAGGTCAAGCCTGGTTCCACTGTTAAGTGTGAGCCTGGATAT 60

Qy     121 gacagccttgtccacatttctcaggcttcacttgactgcacagtgaaatctggagaatct 180
           | |    |  | ||| |||| |||||| | ||||   |   |  |||     | ||   |
Db      61 GGCTTTGTGCTGCACCTTTCCCAGGCTGCTCTTG---GGGAATCGAAGAAGAGTGATAAT 117

Qy     181 gtggttttgagtgtgactgttggtggggctaaacttgttattggaacactttcacaagac 240
           |   |  ||  ||| |   ||| ||      |||||||  |||||||| || ||    |||
Db     118 GCCTTGATGTATGTCAAAATTGATGATCAGAAACTTGCCATTGGAACCCTCTCTGTTGAC 177

Qy     241 aagttccctcagattagctttgatttggtttttgataaagagtttgagctttcacacagc 300
           |||  ||| || ||    |||||| || |||| ||||||||||||||||| || ||||
Db     178 AAGAACCCACACATACAATTTGATCTGATTTTCGATAAAGAGTTTGAGCTGTCGCACACA 237

Qy     301 ggtaccaaagcaaatgttcatttcattggctacaaatccccaacatcgagcaggatgac 360
           |  |   | |  |||   |||| |||||||||         |        |||| ||
Db     238 TCAAAAACTACCAGCGTTTTCTTCACTGGCTACAAGGTTGAGCAGCCATTCGAGGAAGAT 297

Qy     361 ttcactagttcggatgatgaggatgttcctgaagctgttcctgctcctgccctactgct 420
           |    |  |||  ||| ||||     || |   |  || ||          | ||
Db     298 GAAATGGATCTTGATTCTGAAGATGAAGACGAGGAGCTGAATGTTC---------CAGCA 348
```

```
Qy    421 gttactgccaacggaaatgctggagcagctgttgtcaaggctgacacaaagccaaaggcc 480
          || |  |  || || || ||||  |          ||   | |   |||| ||   ||
Db    349 GTCAAGGAAAATGGCAAAGCTGATGAGAAGAAACAGAAAAGTCAAGAAAAGGCAGTTGCT 408

Qy    481 aaacctgccgaagtgaagcctgcagaagagaagcctgaatcagac---------gaggaa 531
          ||||  |  ||   |  || |      |||||    |    |||          || ||
Db    409 GCACCTTCAAAATCAAGTCCGGATTCCAAGAAGAGCAAGGATGACGACGATTCTGATGAG 468

Qy    532 gatgagtctgatgatgaagatgaagtctgaagaggatgatgactcttgagaaaggaatgg 591
          || ||| ||||||||   |||||  |     |  |       |  |   |
Db    469 GACGAGACTGATGATTCTGATGAGGATGAGACGGACGATTCTGATGAGGGTTTGTCTCCT 528

Qy    592 atggttgatgaagatgactcagatgatgacgaggaggaaggattctgaggatgaagaaga 651
            |  |  || ||||| |||  |||||| || ||     |   |   |      |
Db    529 GAAGAAGGCGATGATGATTCAAGTGATGAAGATGATACCAGTGACGATGAGGAGGAAGAC 588

Qy    652 ggaggagctcttaagaagcctgagccaatcaacaagaagaggccaaatgaatctgtatcc 711
            ||| ||||||||||||   |  ||| ||||      |  |    || ||
Db    589 ACTCCAACTCCTAAGAAGCCTGAGGTAGGCAAGAAGAGAGCTGCTGAAAGTTCCGTGCTG 648

Qy    712 aaaacacccgtctctggaaagaaggcaaaaccagcagcagcaccagct 759
          ||||| ||  | ||||  ||||| |||||    ||  || |  || ||
Db    649 AAAACTCCTCTATCTGATAAGAAAGCAAAGGTTGCCACACCGTCATCT 696
```

RESULT 11
AR168371
LOCUS AR168371 1307 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 17 from patent US 6287843.
ACCESSION AR168371
VERSION AR168371.1 GI:17904252
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1307)
AUTHORS Baldwin,D.Adelphi., Briggs,S.P. and Crane,V.C.
TITLE Maize histone deacetylases and their use
JOURNAL Patent: US 6287843-A 17 11-SEP-2001;
FEATURES Location/Qualifiers
source 1. .1307
/organism="unknown"
BASE COUNT 386 a 260 c 308 g 353 t
ORIGIN

Query Match 9.2%; Score 112; DB 6; Length 1307;
Best Local Similarity 54.0%; Pred. No. 6.8e-16;
Matches 282; Conservative 0; Mismatches 225; Indels 15; Gaps 2;

```
Qy     59 caatggagttctggggagttgcggtgacaccaaaaaacgctactaaggtgactcctgaag 118
          ||||||||||||||| |  |  ||||  |||  |  | |   |      |    ||  |
Db     85 CAATGGAGTTCTGGGGTGAAGAAGTGAAGCCAGGAGCCACGGTTTCTTGCAAAGTTGGTG 144

Qy    119 aagacagccttgtccacatttctcaggcttcacttgactgcacagtgaaatctggagaat 178
          | |      || ||||| |||| |||||| | || |      ||  |||| |  | ||
Db    145 ATGGTTTGGTTATCCACCTTTCACAGGCTGCCCTAGGGGAACCAAAGAAAGCGAGTGAGA 204

Qy    179 ctgtggttttgagtgtgactgttggtggggctaaacttgttattggaacactttcacaag 238
           ||   || ||  ||| |   ||| ||      ||||| ||  |||||||  | |||   |
Db    205 ATGCCATTGTGTCTGTCAAAATTGATGATAAGAAACTAGTGCTTGGAACCTTATCAGTTG 264

Qy    239 acaagttccctcagattagctttgatttggtttttgataaaagagtttgagctttcacaca 298
          | |||   ||||| ||    | |||| |||| ||||||||||| |||||| | |||||||
Db    265 AGAAGCATCCTCAAATCTCTTGTGATCTGGTATTTGATAAAGATTTTGAGTTATCACACA 324

Qy    299 gcggtaccaaagcaaatgttcatttcattggctacaaatcccccaacatcgagcagga-- 356
              |  | ||| | ||||   |||  ||||||||| || ||          |   ||
Db    325 ATTCAAAGACAGCTAGTGTTTTCTTCTGTGGCTACAAGTCACCTGTTCCTCTGTTTGAGT 384

Qy    357 -------tgacttcactagttcggatgatgaggatgttcctgaagctgttcctgctcc-- 407
                 ||       |||||| ||||| |||| ||  || ||     ||||    |
```

```
Db    385 CTGATTCTGGTGAAGATAGTTCAGATGAAGAGGTTGAACCCGATCTAATTCCAATGCAGA 444

Qy    408 ----tgcccctactgctgttactgccaacggaaatgctggagcaqctgttgtcaaggctg 463
              ||    ||     |  |||||| || |      |      |  | | |    || ||||
Db    445 ATAATGAAATTAAAATTTCTACTGCAAAGGTTCCCGTGAAGGTTGGTATACAAAATGCTG 504

Qy    464 acacaaagccaaaggccaaacctgccgaagtgaagcctgcagaagagaagcctgaatcag 523
          |   | |   ||  | |    ||  || |  |     | ||     |   ||||
Db    505 ATGAAGATGAAACTTCTAGTGGTGATGATGATGATTTCACTGATAGTGATAGTGAAATGT 564

Qy    524 acgaggaagatgagtctgatgatgaagatgaagtctgaagag 565
            |||||||||||||   |||||||||||||||| | ||| |
Db    565 CTGAGGAAGATGAGTCCAGTGATGAAGATGAAGTGTCAAGCG 606
```

RESULT 12
AR168369
LOCUS AR168369 1191 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 13 from patent US 6287843.
ACCESSION AR168369
VERSION AR168369.1 GI:17904249
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1191)
AUTHORS Baldwin,D.Adelphi., Briggs,S.P. and Crane,V.C.
TITLE Maize histone deacetylases and their use
JOURNAL Patent: US 6287843-A 13 11-SEP-2001;
FEATURES Location/Qualifiers
source 1. .1191
/organism="unknown"
BASE COUNT 365 a 242 c 311 g 273 t
ORIGIN

Query Match 9.2%; Score 111.6; DB 6; Length 1191;
Best Local Similarity 51.4%; Pred. No. 8.4e-16;
Matches 362; Conservative 0; Mismatches 329; Indels 13; Gaps 4;

```
Qy     61 atggagttctggggagttgcggtgacaccaaaaaacgctactaaggtgactcctgaagaa 120
          ||||||||||||||  | | ||| |  ||      | ||  ||||      |||| | |
Db     63 ATGGAGTTCTGGGGTCTCGAGGTCAAGCCTGGTTCCACTGTTAAGTGTGAGCCTGGATAT 122

Qy    121 qacagccttgtccacatttctcaggcttcacttgactgcacagtgaaatctggagaatct 180
          | |     |  | ||| |||| |||||| | ||||   |   |  |||     | ||   |
Db    123 GGCTTTGTGCTGCACCTTTCCCAGGCTGCTCTTG---GGGAATCGAAGAAGAGTGATAAT 179

Qy    181 gtggttttgagtgtgactgttggtggggctaaacttgttattggaacactttcacaagac 240
          |   |  ||  ||| |   ||| ||      |||||||  |||||||| || ||    |||
Db    180 GCCTTGATGTATGTCAAAATTGATGATCAGAAACTTGCCATTGGAACCCTCTCTGTTGAC 239

Qy    241 aagttccctcagattagctttgatttggtttttgataaaagagtttgagctttcacacagc 300
          |||  ||| || |||   |||||| || |||| ||||||||||||||||| || ||||
Db    240 AAGAACCCACACATTCAATTTGATCTGATTTTCGATAAAGAGTTTGAGCTGTCGCACACA 299

Qy    301 ggtaccaaagcaaatgttcatttcattggctacaaatcccccaacatcgagcaggatgac 360
          |  |   | |  ||    |||| |||||||||         |         |||| ||
Db    300 TCAAAAACTACCAGCGTCTTCTTCACTGGCTACAAGGTTGAACAGCCATTCGAGGAAGAT 359

Qy    361 ttcactagttcggatgatgaggatgttcctgaagctgttcctgctcctgcccctactg-- 418
          |    |  ||| ||| ||||     || |   |  || ||| |     |  |
Db    360 GAAATGGATCTTGATTCTGAAGATGAAGACGAGGAGCTGAATGTTCCAGTAGTCAAGGAA 419

Qy    419 -ctgttactgccaacggaaatgctggagcagctgttgtcaaggctgacacaaagccaaag 477
           ||  |  ||  | || ||          |  |   |  ||||| |   |    ||
Db    420 AATGGCAAAGCTGATGGGAAGAAACAGAAAAGTCAAGAAAAGGCAGTTGCTGCACCTTCA 479

Qy    478 gccaaacctgccgaagtgaagcctgcagaagagaagcctgaatcagacgaggaagatgag 537
            |  | | ||   |||       | ||  |   ||| || || |||||  ||
Db    480 AAATCAAGTCCGGATTCCAAGAAGAGCAAGGATGACGATGATTCTGATGAGGACGAGACT 539

Qy    538 tctgatgatgaagatga--agtctgaagaggatgatgactcttgagaaaggaatggatgg 595
```

```
Db    540 GATGATTCTGATGAGGATGAGACGGACGATTCTGATGAGGGTTTATCTTCTGAAGAA--- 596

Qy    596 ttgatgaagatgactcagatgatgacgaggaggaaggattctgaggatgaagaagaggag 655

Db    597 --GGCGATGATGATTCAAGTGATGAAGATGATACCAGTGACGATGAGGAGGAAGACACTC 654

Qy    656 gagctcttaagaagcctgagccaatcaacaagaagaggccaaatgaatctgtatccaaaa 715

Db    655 CAACTCCTAAGAAGCCTGAGGTAGGCAAGAAGAGAGCTGCTGAAAGTTCCGTGCTGAAAA 714

Qy    716 cacccgtctctggaaagaaggcaaaaccagcagcagcaccagct 759

Db    715 CTCCTCTATCTGATAAGAAAGCAAAGGTTGCCACACCGTCATCT 758


RESULT 13
AR168370
LOCUS       AR168370               1245 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 15 from patent US 6287843.
ACCESSION   AR168370
VERSION     AR168370.1  GI:17904250
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1245)
  AUTHORS   Baldwin,D.Adelphi., Briggs,S.P. and Crane,V.C.
  TITLE     Maize histone deacetylases and their use
  JOURNAL   Patent: US 6287843-A 15 11-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .1245
                     /organism="unknown"
BASE COUNT      385 a    250 c    324 g    286 t
ORIGIN


  Query Match             9.1%;  Score 111.2;  DB 6;  Length 1245;
  Best Local Similarity   51.6%;  Pred. No. 1e-15;
  Matches  430;  Conservative    0;  Mismatches  363;  Indels   41;  Gaps    6;

Qy     13 taaaaaaaaacctaacaacctctcttctctcttcctcgttcaacaacaatggagttctgg 72

Db     36 TTAAACCCTAGTTCCCATTCCATCTTCGTTTCCGCTCCGCCGCCGTCGATGGAGTTCTGG 95

Qy     73 ggagttgcggtgacaccaaaaaacgctactaaggtgactcctgaagaagacagccttgtc 132

Db     96 GGTCTCGAGGTCAAACCTGGATCCACTGTCAAGTGTGAGCCTGGACATGGCTTTATCCTG 155

Qy    133 cacatttctcaggcttcacttgactgcacagtgaaatctggagaatctgtggttttgagt 192

Db    156 CACGTTTCCCAGGCTGCCCTTGGGGAATCAAAGAAAAGT---GACAGTGCCTTAATGTAT 212

Qy    193 gtgactgttggtggggctaaacttgttattggaacactttcacaagacaagttccctcag 252

Db    213 GTCAAAGTTGATGACAAGAAGCTTGCCATTGGAACGCTCTCTATCGACAAATACCCACAG 272

Qy    253 attagctttgatttggtttttgataaagagtttgagctttcacacagcggtaccaaagca 312

Db    273 ATACAATTCGATTTGGTTTTCAATAAAGAGTTTGAGCTGTCACACACATCGAAAACTACC 332

Qy    313 aatgttcatttcattggctacaaatcccccaacatcgagcaggatgacttcactagttcg 372

Db    333 AGTGTATTTTTCTCTGGTTACAAGGTTGAGCAGCCAATTGAGGGAGATGAAATGGATCTT 392

Qy    373 gatgatgaggatgttcctgaagctgttcctgctcctg-----------ccctactgct 420

Db    393 GATTCTGAGGATGAAGAGGAGGAGCTAAACATTCCAGTAATCAAGGAAAATGGCAAAGCT 452

Qy    421 gttactgccaacggaaatgctggagcagctgttgtcaaggctgacacaaagccaaaggcc 480

Db    453 GATGGGAAGGAGGAGCAGAAAAATCAAGAGAAGGCAGTAGCTGCTACAGCTTCAAAATCA 512

Qy    481 aaacctgccgaagtgaagcctgcagaagagaagcctgaatcagacgaggaaagatgagtct 540

Db    513 AGTCTTGGCCTTGAAAAGAAAAGCAAGGATGACTCTGATGATTCTGATGAGGATGAGTCT 572

Qy    541 gatgatgaagatgaagtctgaagaggatgatgactcttgagaaaggaatggat------- 593

Db    573 GATGATTCTGATGA----GGATGATTCTGATGATTCTGATGAAGGCGAGGGATTATCTCC 628

Qy    594 ---ggttgatgaagatgactcagatgatgacgaggaggaaggattctgaggatgaagaag 650

Db    629 TGACGAAGGCGATGATGATTCAAGTGATGAGGATGATACCAGTGATGATGACGAGGAAGA 688

Qy    651 aggaggagctcttaagaagcctgagccaatcaacaagaagaggccaaatgaatctgtatc 710

Db    689 AACCCCAACTCCTAAAAAGCCAGAGGCA---GGCAAGAAGAGAGGTGCTGAAAATGCTCT 745

Qy    711 caaaacacccgtctctggaaagaaggcaaaaccagcagcagcaccagcttctactcctca 770

Db    746 GAAAACGCCTCTTTCTGATAAGAAAGCAAAGGTTGCCACACCGCCAG---------CCCA 796

Qy    771 gaagacagaagagaagaagaaaggaggacacaccgccacaccacacccagctaa 824

Db    797 GAAAACAGGTGGCAAGAAGGGTGCCACCCATGTGGCAACTCCACACCCAGCAAA 850


RESULT 14
AR168368
LOCUS       AR168368               1283 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 11 from patent US 6287843.
ACCESSION   AR168368
VERSION     AR168368.1  GI:17904247
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1283)
  AUTHORS   Baldwin,D.Adelphi., Briggs,S.P. and Crane,V.C.
  TITLE     Maize histone deacetylases and their use
  JOURNAL   Patent: US 6287843-A 11 11-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .1283
                     /organism="unknown"
BASE COUNT      383 a    259 c    333 g    308 t
ORIGIN


  Query Match             9.1%;  Score 111.2;  DB 6;  Length 1283;
  Best Local Similarity   50.7%;  Pred. No. 1e-15;
  Matches  359;  Conservative    0;  Mismatches  328;  Indels   21;  Gaps    3;

Qy     61 atggagttctggggagttgcggtgacaccaaaaaacgctactaaggtgactcctgaagaa 120

Db     79 ATGGAGTTCTGGGGTCTCGAGGTCAAGCCTGGTTCCACTGTTAAGTGTGAGCCTGGATAT 138

Qy    121 gacagccttgtccacatttctcaggcttcacttgactgcacagtgaaatctggagaatct 180

Db    139 GGCTTTGTGCTGCACCTTTCCCAGGCTGCTCTTG---GGGAATCGAAGAAGAGTGATAAT 195

Qy    181 gtggttttgagtgtgactgttggtggggctaaacttgttattggaacactttcacaagac 240

Db    196 GCCTTGATGTATGTCAAAATTGATGATCAGAAACTTGCCATTGGAACCCTCTCTGTTGAC 255

Qy    241 aagttccctcagattagctttgatttggtttttgataaagagtttgagctttcacacagc 300

Db    256 AAGAACCCACACATACAATTTGATCTGATTTTCGATAAAGAGTTTGAGCTGTCGCACACA 315

Qy    301 ggtaccaaagcaaatgttcatttcattggctacaaatccccccaacatcgaqcaggatgac 360

Db    316 TCAAAAACTACCAGCGTTTTCTTCACTGGCTACAAGGTTGAGCAGCCATTCGAGGAAGAT 375

Qy    361 ttcactagttcggatgatgaggatgttcctgaagctgttcctgctcctgccccctactgct 420

Db    376 GAAATGGATCTTGATTCTGAAGATGAAGACGAGGAGCTGAATGTTC---------CAGTA 426
```

```
Qy    421 gttactgccaacggaaatgctggagcagctgttgtcaaggctgacacaaagccaaaggcc 480
          || |  |  || || || ||||  |           ||   | |   |||| ||   ||
Db    427 GTCAAGGAAAATGGCAAAGCTGATGAGAAGAAACAGAAAAGTCAAGAAAAGGCAGTTGCT 486

Qy    481 aaacctgccgaagtgaagcctgcagaagagaagcctgaatcagac---------gaggaa 531
            |||| |  ||   |  || |      |||||    |    |||         || ||
Db    487 GCACCTTCAAAATCAAGTCCGGATTCCAAGAAGAGCAAGGATGACGACGATTCTGATGAG 546

Qy    532 gatgagtctgatgatgaagatgaagtctgaagaggatgatgactcttgagaaaggaatgg 591
          || ||| ||||||||   ||||| |     |  |     |       | |   |
Db    547 GACGAGACTGATGATTCTGATGAGGATGAGACGGACGATTCTGATGAGGGTTTGTCTCCT 606

Qy    592 atggttgatgaagatgactcagatgatgacgaggaggaaggattctgaggatgaagaaga 651
            |  |  || ||||| |||  |||||| || ||     |   |   |         |
Db    607 GAAGAAGGCGATGATGATTCAAGTGATGAAGATGATACCAGTGACGATGAGGAGGAAGAC 666

Qy    652 ggaggagctcttaagaagcctgagccaatcaacaagaagaggccaaatgaatctgtatcc 711
            ||| ||||||||||||||  |  ||| ||||      |  |    || ||
Db    667 ACTCCAACTCCTAAGAAGCCTGAGGTAGGCAAGAAGAGAGCTGCTGAAAGTTCCGTGCTG 726

Qy    712 aaaacacccgtctctggaaagaaggcaaaaccagcagcagcaccagct 759
          |||||  ||  | ||||  ||||| |||||    ||  || |  || ||
Db    727 AAAACTCCTCTATCTGATAAGAAAGCAAAGGTTGCCACACCGTCATCT 774


RESULT  15
ZMU82815
LOCUS       ZMU82815                1169 bp    mRNA    linear   PLN 17-JUL-1997
DEFINITION  Zea mays nucleolar histone deacetylase HD2-p39 mRNA, complete cds.
ACCESSION   U82815
VERSION     U82815.1  GI:2257755
KEYWORDS    .
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 1169)
  AUTHORS   Lusser,A., Brosch,G., Loidl,A., Haas,H. and Loidl,P.
  TITLE     Identification of maize histone deacetylase HD2 as an acidic
            nucleolar phosphoprotein
  JOURNAL   Science 277 (5322), 88-91 (1997)
  MEDLINE   97349336
REFERENCE   2  (bases 1 to 1169)
  AUTHORS   Lusser,A., Brosch,G., Loidl,A., Haas,H. and Loidl,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JAN-1997) Department of Microbiology, University of
            Innsbruck, Fritz Preql-Strasse 3, Innsbruck 6020, Austria
FEATURES             Location/Qualifiers
     source          1. .1169
                     /organism="Zea mays"
                     /strain="Cuzco 251"
                     /db_xref="taxon:4577"
     CDS             74. .997
                     /codon_start=1
                     /product="nucleolar histone deacetylase HD2-p39"
                     /protein_id="AAB63262.1"
                     /db_xref="GI:2257756"
                     /translation="MEFWGLEVKPGSTVKCEPGYGFVLHLSQAALGESKKSDNALMYV
                     KIDDQKLAIGTLSVDKNPHIQFDLIFDKEFELSHTSKTTSVFFTGYKVEQPFEEDEMD
                     LDSEDEDEELNVPVVKENGKADEKKQKSQEKAVAAPSKSSPDSKKSKDDDDSDEDETD
                     DSDEDETDDSDEGLSSEEGDDDSSDEDDTSDDEEEDTPTPKKPEVGKKRPAESSVLKT
                     PLSDKKAKVATPSSQKTGGKKGAAVHVATPHPAKGKTIVNNDKSVKSPKSAPKSGGSV
                     PCKPCSKSFISFTALQAHSRAKMGASESQVQ"
BASE COUNT      344 a    246 c    310 g    269 t
ORIGIN


  Query Match             9.0%;  Score 110;  DB 8;  Length 1169;
  Best Local Similarity   51.3%;  Pred. No. 2e-15;
  Matches  361;  Conservative    0;  Mismatches  330;  Indels   13;  Gaps    4;

Qy     61 atggagttctggggtctgaggtcaagcctggttccactgttaagtgtgagcctggatat 120
          |||||||||||||||  | | ||| |  ||      | ||  ||||      |||| | |
Db     74 ATGGAGTTCTGGGGTCTCGAGGTCAAGCCTGGTTCCACTGTTAAGTGTGAGCCTGGATAT 133

Qy    121 gacagccttgtccacatttctcaggcttcacttgactgcacagtgaaatctggagaatct 180
          | |    |  | ||| |||| |||||| | ||||   |   |  |||     | ||   |
Db    134 GGCTTTGTGCTGCACCTTTCCCAGGCTGCTCTTG---GGGAATCGAAGAAGAGTGATAAT 190

Qy    181 gtggttttgagtgtgactgttggtggggctaaacttgttattggaacactttcacaagac 240
          |   |  ||  ||| |   ||| ||     |||||||  ||||||| || ||    |||
Db    191 GCCTTGATGTATGTCAAAATTGATGATCAGAAACTTGCCATTGGAACCCTCTCTGTTGAC 250

Qy    241 aagttccctcagattagctttgatttggtttttgataaagagtttgagctttcacacagc 300
          |||  ||| || |||   |||||| || |||| ||||||||||||||||| || ||||
Db    251 AAGAACCCACACATTCAATTTGATCTGATTTTCGATAAAGAGTTTGAGCTGTCGCACACA 310

Qy    301 ggtaccaaagcaaatgttcatttcattggctacaaatcccccaacatcgagcaggatgac 360
            |  |  | |  ||    |||| |||||||||       |       |||| ||
Db    311 TCAAAAACTACCAGCGTCTTCTTCACTGGCTACAAGGTTGAACAGCCATTCGAGGAAGAT 370

Qy    361 ttcactagttcggatgatgaggatgttcctgaagctgttcctgctcctgcccctactg-- 418
            |   |  |||  ||| ||||    || |  |  || ||| |     |  |
Db    371 GAAATGGATCTTGATTCTGAAGATGAAGACGAGGAGCTGAATGTTCCAGTAGTCAAGGAA 430

Qy    419 -ctgttactgccaacggaaatgctggagcagctgttgtcaaggctgacacaaagccaaag 477
           ||  |  ||  | |  ||         |  |   |  |||||  |   |   ||
Db    431 AATGGCAAGGCTGATGAGAAGAAACAGAAAAGTCAAGAAAAGGCAGTTGCTGCACCCTCA 490

Qy    478 gccaaacctgccgaagtgaagcctgcagaagagaagcctgaatcagacgaggaagatgag 537
            |  | |  ||   |||       | ||  |  |||  || || ||||| ||
Db    491 AAATCAAGTCCGGATTCCAAGAAGAGCAAGGATGACGATGATTCTGATGAGGACGAGACT 550

Qy    538 tctgatgatgaagatga--agtctgaagaggatgatgactcttgagaaaggaatggatgg 595
            ||||  ||| || ||  || | || ||   ||||||   || |       | | |
Db    551 GATGATTCTGATGAGGATGAGACGGACGATTCTGATGAGGGTTTATCTTCTGAAGAA--- 607

Qy    596 ttgatgaagatgactcagatgatgacgaggaggaaggattctgaggatgaagaagaggag 655
            |  || ||||| |||  |||||| || ||      |   |   |       |
Db    608 --GGCGATGATGATTCAAGTGATGAAGATGATACCAGTGACGATGAGGAGGAAGACACTC 665

Qy    656 gagctcttaagaagcctgagccaatcaacaagaagaggccaaatgaatctgtatccaaaa 715
            | ||| ||||||||||||||  |  ||| ||||      |  |    || ||    ||||
Db    666 CAACTCCTAAGAAGCCTGAGGTAGGCAAGAAGAGACCTGCTGAAAGTTCCGTGCTGAAAA 725

Qy    716 cacccgtctctggaaagaaggcaaaaccagcagcagcaccagct 759
          | ||  | ||||  ||||| |||||    ||  || |  || ||
Db    726 CTCCTCTATCTGATAAGAAAGCAAAGGTTGCCACACCGTCATCT 769


Search completed: May  2, 2002, 20:14:50
Job time: 15504 sec
```

```
            Gorgone,G., Burns,D., Griffin,J., Mouanoutoua,M., Nguyen,D., Tan,R.
            , Rose,M., Warren,B., Ton,B., Kastury,K., Borillo,C., Carpio,T.,
            Policky,J., Suzuki,G., Argentine,C., Shah,S., Nobriga,A., Murry,L.,
            Turner,C., Krikorian,S., Elder,L. and Hanson,D.
  TITLE     Arabidopsis thaliana Gene Expression MicroArray
  JOURNAL   Unpublished (1999)
COMMENT     Contact: David Smoller, Ph.D.
            Genome Systems, Inc., a wholly owned subsidiary of Incyte
            Pharmaceuticals, Inc.
            4633 World Parkway Circle, St. Louis, MO 63134, USA
            Tel: 877-577-2733
            Fax: 314-427-3324
            Email: service@genomesystems.com.
FEATURES             Location/Qualifiers
     source          1. .585
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="701931423"
                     /clone_lib="A. thaliana, mixed source"
                     /note="This sequence was obtained from a clone generated
                     with a PCR product of the target gene."
BASE COUNT       95 a    141 c    132 g    192 t     25 others
ORIGIN


  Query Match             37.6%;  Score 458.4;  DB 9;  Length 585;
  Best Local Similarity   95.3%;  Pred. No. 9.3e-55;
  Matches  525;  Conservative    0;  Mismatches   20;  Indels    6;  Gaps    6;

Qy     298 agcggtaccaaagcaaatgttcatttcattggctacaaatcccccaacatcgagcaggat 357
           |||| ||||||||| |||   |||| ||||||||||||||||||||||||||||||| ||
Db     546 AGCGNTACCAAAGC-AATNNNCATTNCATTGGCTACAAATCCCCCAACATCGAGCAGNAT 488

Qy     358 gacttcactagttcggatgatgaggatgttcctgaagctgttcctgctcctgcccctact 417
           ||||||||||||||| |||||||| |||||||| |||||||||||||||||||||| |||
Db     487 GACTTCACTAGTTCGNATGATGAGNATGTTCCTNAAGCTGTTCCTGCTCCTGCCCCNACT 428

Qy     418 gctgttactgccaacggaaatgctggagcagctgttgtcaaggctgacacaaagccaaag 477
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     427 GCTGTTACTGCCAACGGAAATGCTGGAGCAGCTGTTGTCAAGGCTGACACAAAGCCAAAG 368

Qy     478 gccaaacctgccgaagtgaagcctgcagaagagaagcctgaatcagacgaggaagatgag 537
           |||||||||||||||| |||||||||||||||| ||| ||||||||||!||||||||
Db     367 GCCAAACCTGCCGAAGTNAAGCCTGCAGAAGAGAANCCTNAATCAGACGAGGAAGATGAN 308

Qy     538 tctgatgatgaagatgaagtctgaagaggatgatgactcttgagaaaggaatggatggtt 597
           |||||||||||||||| | |||||||||||||||||||| |||||||||||||||| |||
Db     307 TCTGATGATGAAGATG-ANTCTGAAGAGGATGATGACTC-TGAGAAAGGAATGGAT-GTT 251

Qy     598 gatgaagatgactcagatgatgacgaggaggaaggattctgaggatgaagaagaggagga 657
           ||||||||| |||||||||||||||||||| |||||||||||||||||||||||||||||
Db     250 GATGAAGATNACTCAGATGATGACGAGGAGG-AGGATTCTGAGGATGAAGAAGAGGAGGA 192

Qy     658 g-ctcttaagaagcctgagccaatcaacaagaagaggccaaatgaatctgtatccaaaac 716
           | ||| ||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     191 GACTCCTAAGAAGCCTGAGCCAATCAACAAGAAGAGGCCAAATGAANCTGTATCCAAAAC 132

Qy     717 acccgtctctggaaagaaggcaaaaccagcagcagcaccagcttctactcctcagaagac 776
           |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db     131 ACCCGTCTCTGGAAAGAAGGCAAAACCAGCAGCAGCACCAGNTTCTACTCCTCAGAAGAC 72

Qy     777 agaagagaagaagaaaggaggacacaccgccacaccacacccagctaagaagggtggaaa 836
           ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db      71 AGAAGAGAAGAAGAAAGGAGGACACACCGCCACACCACACNCAGCTAAGAAGGGTGGAAA 12

Qy     837 gtctcctgtga 847
           ||||||||||||
Db      11 GTCTCCTGTGA 1


RESULT   9
AV539421/c
LOCUS       AV539421                 426 bp    mRNA     linear   EST 06-SEP-2000
DEFINITION  AV539421 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
            cDNA clone RZ131f01F 3', mRNA sequence.
ACCESSION   AV539421
VERSION     AV539421.1  GI:8701178
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 426)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     A large scale analysis of cDNA in Arabidopsis thaliana: Generation
            of 12,028 non-redundant expressed sequence tags from normalized and
            size-selected cDNA libraries
  JOURNAL   DNA Res. 7, 175-180 (2000)
  MEDLINE   20363093
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .426
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /clone="RZ131f01F"
                     /clone_lib="Arabidopsis thaliana roots Columbia"
                     /tissue_type="roots"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT      111 a    103 c     74 g    138 t
ORIGIN


  Query Match             35.0%;  Score 426;  DB 9;  Length 426;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-50;
  Matches  426;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     755 cagcttctactcctcagaagacagaagagaagaagaaagaaggaggacacaccgccacaccac 814
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     426 CAGCTTCTACTCCTCAGAAGACAGAAGAGAAGAAGAAAGGAGGACACACCGCCACACCAC 367

Qy     815 acccagctaagaagggtggaaagtctcctgtgaatgctaaccagagccccaagtctggag 874
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     366 ACCCAGCTAAGAAGGGTGGAAAGTCTCCTGTGAATGCTAACCAGAGCCCCAAGTCTGGAG 307

Qy     875 gtcaatcatccggtggtaacaacaacaagaagccattcaactcaggcaaacaatttggtg 934
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     306 GTCAATCATCCGGTGGTAACAACAACAAGAAGCCATTCAACTCAGGCAAACAATTTGGTG 247

Qy     935 gttccaacaacaagggttctaacaagggcaagggaaagggtagagcttaaggacgtggat 994
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     246 GTTCCAACAACAAGGGTTCTAACAAGGGCAAGGGAAAGGGTAGAGCTTAAGGACGTGGAT 187

Qy     995 caaggagaggttttgggttttcgagtagatgatgaaaacacttggaagtgtggttttgga 1054
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     186 CAAGGAGAGGTTTTGGGTTTTCGAGTAGATGATGAAAACACTTGGAAGTGTGGTTTTGGA 127

Qy    1055 tttttatcttatttttattagtataacttgttatcggatgagctattttgagtatttgcaa 1114
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     126 TTTTTATCTTATTTTATTAGTATAACTTGTTATCGGATGAGCTATTTTGAGTATTTGCAA 67

Qy    1115 tttctactttcctatgtaattcagtatatgaatatttgctgaaatgagaaagaagactcg 1174
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66 TTTCTACTTTCCTATGTAATTCAGTATATGAATATTTGCTGAAATGAGAAAGAAGACTCG 7

Qy    1175 aattgc 1180
           ||||||
Db       6 AATTGC 1
```

```
     source          1. .483
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /clone="SQ114b02F"
                     /clone_lib="Arabidopsis thaliana green siliques Columbia"
                     /tissue_type="green siliques"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT      120 a    114 c     89 g    160 t
ORIGIN


  Query Match             39.4%;  Score 479.8;  DB 9;  Length 483;
  Best Local Similarity   99.6%;  Pred. No. 1.1e-57;
  Matches  481;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     697 aatgaatctgtatccaaaacacccgtctctggaaagaaggcaaaaccagcagcagcacca 756
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     483 AATGAATCTGTATCCAAAACACCCGTCTCTGGAAAGAAGGCAAAACCAGCAGCAGCACCA 424

Qy     757 gcttctactcctcagaagacagaagagaagaagaaaggaggacacaccgccacaccacac 816
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db     423 GCTTCTACTCCTCAGAAGCCAGAAGAGAAGAAGAAAGGAGGACACACCGCCACACCACAC 364

Qy     817 ccagctaagaagggtggaaagtctcctgtgaatgctaaccagagccccaagtctggaggt 876
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     363 CCAGTTAAGAAGGGTGGAAAGTCTCCTGTGAATGCTAACCAGAGCCCCAAGTCTGGAGGT 304

Qy     877 caatcatccggtggtaacaacaacaagaagccattcaactcaggcaaacaatttggtggt 936
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 CAATCATCCGGTGGTAACAACAACAAGAAGCCATTCAACTCAGGCAAACAATTTGGTGGT 244

Qy     937 tccaacaacaagggttctaacaagggcaagggaaagggtagagcttaaggacgtggatca 996
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 TCCAACAACAAGGGTTCTAACAAGGGCAAGGGAAAGGGTAGAGCTTAAGGACGTGGATCA 184

Qy     997 aggagaggttttgggttttcgagtagatgatgaaaacacttggaagtgtggttttggatt 1056
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 AGGAGAGGTTTTGGGTTTTCGAGTAGATGATGAAAACACTTGGAAGTGTGGTTTTGGATT 124

Qy    1057 tttatcttattttattagtataacttgttatcggatgagctattttgagtatttgcaatt 1116
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123 TTTATCTTATTTTATTAGTATAACTTGTTATCGGATGAGCTATTTTGAGTATTTGCAATT 64

Qy    1117 tctactttcctatgtaattcagtatatgaatatttgctgaaatgagaaagaagactcgaa 1176
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63 TCTACTTTCCTATGTAATTCAGTATATGAATATTTGCTGAAATGAGAAAGAAGACTCGAA 4

Qy    1177 ttg 1179
           |||
Db       3 TTG 1


RESULT   7
AV565954/c
LOCUS       AV565954                477 bp    mRNA    linear   EST 07-SEP-2000
DEFINITION  AV565954 Arabidopsis thaliana green siliques Columbia Arabidopsis
            thaliana cDNA clone SQ233g03F 3', mRNA sequence.
ACCESSION   AV565954
VERSION     AV565954.1  GI:8737404
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 477)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     A large scale analysis of cDNA in Arabidopsis thaliana: Generation
            of 12,028 non-redundant expressed sequence tags from normalized and
            size-selected cDNA libraries
  JOURNAL   DNA Res. 7, 175-180 (2000)
  MEDLINE   20363093
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES              Location/Qualifiers
     source           1. .477
                      /organism="Arabidopsis thaliana"
                      /strain="Columbia"
                      /db_xref="taxon:3702"
                      /clone="SQ233g03F"
                      /clone_lib="Arabidopsis thaliana green siliques Columbia"
                      /tissue_type="green siliques"
                      /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                      XhoI"
BASE COUNT      117 a    113 c     90 g    157 t
ORIGIN


  Query Match             39.0%;  Score 475.4;  DB 9;  Length 477;
  Best Local Similarity   99.8%;  Pred. No. 4.4e-57;
  Matches  476;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     705 tgtatccaaaacacccgtctctggaaagaaggcaaaaccagcagcagcaccagcttctac 764
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     477 TGTATCCAAAACACCCGTCTCTGGAAAGAAGGCAAAACCAGCAGCAGCACCAGCTTCTAC 418

Qy     765 tcctcagaagacagaagagaagaagaaaggaggacacaccgccacaccacacccagctaa 824
           |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db     417 TCCTCAGAAGCCAGAAGAGAAGAAGAAAGGAGGACACACCGCCACACCACACCCAGCTAA 358

Qy     825 gaagggtggaaagtctcctgtgaatgctaaccagagccccaagtctggaggtcaatcatc 884
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     357 GAAGGGTGGAAAGTCTCCTGTGAATGCTAACCAGAGCCCCAAGTCTGGAGGTCAATCATC 298

Qy     885 cggtggtaacaacaacaagaagccattcaactcaggcaaacaatttggtggttccaacaa 944
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     297 CGGTGGTAACAACAACAAGAAGCCATTCAACTCAGGCAAACAATTTGGTGGTTCCAACAA 238

Qy     945 caagggttctaacaagggcaagggaaagggtagagcttaaggacgtggatcaaggagagg 1004
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     237 CAAGGGTTCTAACAAGGGCAAGGGAAAGGGTAGAGCTTAAGGACGTGGATCAAGGAGAGG 178

Qy    1005 ttttgggttttcgagtagatgatgaaaacacttggaagtgtggttttggatttttatctt 1064
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     177 TTTTGGGTTTTCGAGTAGATGATGAAAACACTTGGAAGTGTGGTTTTGGATTTTTATCTT 118

Qy    1065 attttattagtataacttgttatcggatgagctattttgagtatttgcaatttctacttt 1124
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117 ATTTTATTAGTATAACTTGTTATCGGATGAGCTATTTTGAGTATTTGCAATTTCTACTTT 58

Qy    1125 cctatgtaattcagtatatgaatatttgctgaaatgagaaagaagactcgaattgca 1181
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57 CCTATGTAATTCAGTATATGAATATTTGCTGAAATGAGAAAGAAGACTCGAATTGCA 1


RESULT   8
AW004431/c
LOCUS       AW004431                585 bp    mRNA    linear   EST 08-SEP-1999
DEFINITION  701931423 A. thaliana, mixed source Arabidopsis thaliana cDNA clone
            701931423, mRNA sequence.
ACCESSION   AW004431
VERSION     AW004431.1  GI:5851460
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 585)
  AUTHORS   Chen,J., Momiyama,M., Chan,E., Mooney,M., Carroon,B., Gilliland,D.,
            Wang,X., Hillman,J., Guegler,K., Kim,C., Doyle,M., Brzoska,P.,
```

```
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db    552 GATGAAGATGACTCAGATGATGACGAGGAGG-AGGATTCTGAGGATGAAGAAGAGGAGGA 494

Qy    658 g-ctcttaagaagcctgagccaatcaacaagaagaggccaaatgaatctgtatccaaaac 716
          | ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    493 GACTCCTAAGAAGCCTGAGCCAATCAACAAGAAGAGGCCAAATGAATCTGTATCCAAAAC 434

Qy    717 acccgtctctggaaagaaggcaaaaccagcagcagcaccagcttctactcctcagaagac 776
          ||||||||||||||||||                  ||||||||||||||||||
Db    433 ACCCGTCTCTGGAAAGAAGNNNNNNNNNNNNNNNNNNNNNNTTCTACTCCTCAGAAGAC 374

Qy    777 agaagagaagaagaaaggaggacacaccgccacaccacacccagctaagaagggtggaaa 836
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    373 AGAAGAGAAGAAGAAAGGAGGACACACCGCCACACCACACCCAGCTAAGAAGGGTGGAAA 314

Qy    837 gtctcctgtgaatgctaaccagagccccaagtctggaggtcaatcatccggtggtaacaa 896
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    313 GTCTCCTGTGAATGCTAACCAGAGCCCCAAGTCTGGAGGTCAATCATCCGGTGGTAACAA 254

Qy    897 caacaagaagccattcaactcaggcaaacaatttggtggttccaacaacaagggttctaa 956
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    253 CAACAAGAAGCCATTCAACTCAGGCAAACAATTTGGTGGTTCCAACAACAAGGGTTCTAA 194

Qy    957 caagggcaagggaaagggtagagcttaaggacgtggatcaaggagaggttttgggttttc 1016
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    193 CAAGGGCAAGGGAAAGGGTAGAGCTTAAGGACGTGGATCAAGGAGAGGTTTTGGGTTTTC 134

Qy   1017 gagtagatgatgaaaacacttggaagtgtggttttggatttttatcttatttttattagta 1076
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    133 GAGTAGATGATGAAAACACTTGGAAGTGTGGTTTTGGATTTTTATCTTATTTTATTAGTA 74

Qy   1077 taacttgttatcggatgagctattttgagtatttgcaatttctactttcctatgtaattc 1136
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     73 TAACTTGTTATCGGATGAGCTATTTTGAGTATTTGCAATTTCTACTTTCCTATGTAATTC 14

Qy   1137 agtatatgaatat 1149
          |||||||||||||
Db     13 AGTATATGAATAT 1
```

RESULT 5
AV542413/c
LOCUS AV542413 503 bp mRNA linear EST 07-SEP-2000
DEFINITION AV542413 Arabidopsis thaliana roots Columbia Arabidopsis thaliana cDNA clone RZ183f11F 3', mRNA sequence.
ACCESSION AV542413
VERSION AV542413.1 GI:8704173
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 503)
AUTHORS Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
JOURNAL DNA Res. 7, 175-180 (2000)
MEDLINE 20363093
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES Location/Qualifiers
source 1. .503
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="RZ183f11F"
/clone_lib="Arabidopsis thaliana roots Columbia"
/tissue_type="roots"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 119 a 118 c 95 g 171 t
ORIGIN

```
  Query Match            41.2%;  Score 502;  DB 9;  Length 503;
  Best Local Similarity  100.0%;  Pred. No. 8.7e-61;
  Matches  502;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    684 caagaagaggccaaatgaatctgtatccaaaacacccgtctctggaaagaaggcaaaacc 743
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    503 CAAGAAGAGGCCAAATGAATCTGTATCCAAAACACCCGTCTCTGGAAAGAAGGCAAAACC 444

Qy    744 agcagcagcaccagcttctactcctcagaagacagaagagaagaagaaagaggaggacacac 803
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    443 AGCAGCAGCACCAGCTTCTACTCCTCAGAAGACAGAAGAGAAGAAGAAAGGAGGACACAC 384

Qy    804 cgccacaccacacccagctaagaagggtggaaagtctcctgtgaatgctaaccagagccc 863
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    383 CGCCACACCACACCCAGCTAAGAAGGGTGGAAAGTCTCCTGTGAATGCTAACCAGAGCCC 324

Qy    864 caagtctggaggtcaatcatccggtggtaacaacaacaagaagccattcaactcaggcaa 923
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    323 CAAGTCTGGAGGTCAATCATCCGGTGGTAACAACAACAAGAAGCCATTCAACTCAGGCAA 264

Qy    924 acaatttggtggttccaacaacaagggttctaacaagggcaagggaaagggtagagctta 983
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    263 ACAATTTGGTGGTTCCAACAACAAGGGTTCTAACAAGGGCAAGGGAAAGGGTAGAGCTTA 204

Qy    984 aggacgtggatcaaggagaggttttgggttttcgagtagatgatgaaaacacttggaagt 1043
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    203 AGGACGTGGATCAAGGAGAGGTTTTGGGTTTTCGAGTAGATGATGAAAACACTTGGAAGT 144

Qy   1044 gtggttttggatttttatcttatttttattagtataacttgttatcggatgagctattttg 1103
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    143 GTGGTTTTGGATTTTTATCTTATTTTATTAGTATAACTTGTTATCGGATGAGCTATTTTG 84

Qy   1104 agtatttgcaatttctactttcctatgtaattcagtatatgaatatttgctgaaatgaga 1163
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     83 AGTATTTGCAATTTCTACTTTCCTATGTAATTCAGTATATGAATATTTGCTGAAATGAGA 24

Qy   1164 aagaagactcgaattgcaaaca 1185
          |||||||||||||||||||||||
Db     23 AAGAAGACTCGAATTGCAAACA 2
```

RESULT 6
AV559263/c
LOCUS AV559263 483 bp mRNA linear EST 07-SEP-2000
DEFINITION AV559263 Arabidopsis thaliana green siliques Columbia Arabidopsis thaliana cDNA clone SQ114b02F 3', mRNA sequence.
ACCESSION AV559263
VERSION AV559263.1 GI:8730689
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 483)
AUTHORS Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
JOURNAL DNA Res. 7, 175-180 (2000)
MEDLINE 20363093
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES Location/Qualifiers

```
Qy   1142 atgaatatttgctgaaatgagaaagaagactcgaattgcaaac 1184
          |||||||||||||||||||||||||||||||||||||||||||
Db     45 ATGAATATTTGCTGAAATGAGAAAGAAGACTCGAATTGCAAAC 3


RESULT   3
AV543549/c
LOCUS       AV543549                559 bp    mRNA    linear   EST 07-SEP-2000
DEFINITION  AV543549 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
            cDNA clone RZ202e12F 3', mRNA sequence.
ACCESSION   AV543549
VERSION     AV543549.1  GI:8714963
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 559)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     A large scale analysis of cDNA in Arabidopsis thaliana: Generation
            of 12,028 non-redundant expressed sequence tags from normalized and
            size-selected cDNA libraries
  JOURNAL   DNA Res. 7, 175-180 (2000)
  MEDLINE   20363093
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .559
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /clone="RZ202e12F"
                     /clone_lib="Arabidopsis thaliana roots Columbia"
                     /tissue_type="roots"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT      129 a    132 c    103 g    195 t
ORIGIN


  Query Match             43.9%;  Score 534.8;  DB 9;  Length 559;
  Best Local Similarity   98.6%;  Pred. No. 2.3e-65;
  Matches  550;  Conservative    0;  Mismatches    7;  Indels    1;  Gaps    1;

Qy    628 gaaggattctgaggatgaagaagaggaggag-ctcttaagaagcctgagccaatcaacaa 686
          | |||||||||| || ||||||||||||| ||| |||||||||| ||||||||||||
Db    558 GGAGGATTCTGAGTATTAAGAAGAGGAGGAGACTCCTAAGAAGCCTAAGCCAATCAACAA 499

Qy    687 gaagaggccaaatgaatctgtatccaaaacacccgtctctggaaagaaggcaaaaccagc 746
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    498 AAAGAGGCCAAATGAATCTGTATCCAAAACACCCGTCTCTGGAAAGAAGGCAAAACCAGC 439

Qy    747 agcagcaccagcttctactcctcagaagacagaagagaagaagaaaggaggacacaccgc 806
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    438 AGCAGCACCAGCTTCTACTCCTCAGAAGACAGAAGAGAAGAAGAAAGGAGGACACACCGC 379

Qy    807 cacaccacacccagctaagaagggtggaaagtctcctgtgaatgctaaccagagccccaa 866
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    378 CACACCACACCCAGCTAAGAAGGGTGGAAAGTCTCCTGTGAATGCTAACCAGAGCCCCAA 319

Qy    867 gtctggaggtcaatcatccggtggtaacaacaacaagaagccattcaactcaggcaaaca 926
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    318 GTCTGGAGGTCAATCATCCGGTGGTAACAACAACAAGAAGCCATTCAACTCAGGCAAACA 259

Qy    927 atttggtggttccaacaacaaqqgttctaacaagggcaagggaaagggtagagcttaagg 986
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    258 ATTTGGTGGTTCCAACAACAAGGGTTCTAACAAGGGCAAGGGAAAGGGTAGAGCTTAAGG 199

Qy    987 acgtggatcaaggagaggttttgggttttcgagtagatgatgaaaacacttggaagtgtg 1046
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    198 ACGTGGATCAAGGAGAGGTTTTGGGTTTTCGAGTAGATGATGAAAACACTTGGAAGTGTG 139

Qy   1047 gttttggatttttatcttatttttattagtataacttgttatcggatgagctatttttgagt 1106
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    138 GTTTTGGATTTTTATCTTATTTTTATTAGTATAACTTGTTATCGGATGAGCTATTTTGAGT 79

Qy   1107 atttgcaatttctactttcctatgtaattcagtatatgaatatttgctgaaatgagaaag 1166
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     78 ATTTGCAATTTCTACTTTCCTATGTAATTCAGTATATGAATATTTGCTGAAATGAGAAAA 19

Qy   1167 aagactcgaattgcaaac 1184
          ||||||||||||||||||
Db     18 AAGACTCGAATTGCAAAC 1


RESULT   4
AI998449/c
LOCUS       AI998449                552 bp    mRNA    linear   EST 08-SEP-1999
DEFINITION  701545727 A. thaliana, Columbia Col-0, rosette-2 Arabidopsis
            thaliana cDNA clone 701545727, mRNA sequence.
ACCESSION   AI998449
VERSION     AI998449.1  GI:5845354
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 552)
  AUTHORS   Chen,J., Momiyama,M., Chan,E., Mooney,M., Carroon,B., Gilliland,D.,
            Wang,X., Hillman,J., Guegler,K., Kim,C., Doyle,M., Brzoska,P.,
            Gorgone,G., Burns,D., Griffin,J., Mouanoutoua,M., Nguyen,D., Tan,R.
            , Rose,M., Warren,B., Ton,B., Kastury,K., Borillo,C., Carpio,T.,
            Policky,J., Suzuki,G., Argentine,C., Shah,S., Nobriga,A., Murry,L.,
            Turner,C., Krikorian,S., Elder,L. and Hanson,D.
  TITLE     Arabidopsis thaliana Gene Expression MicroArray
  JOURNAL   Unpublished (1999)
COMMENT     Contact: David Smoller, Ph.D.
            Genome Systems, Inc., a wholly owned subsidiary of Incyte
            Pharmaceuticals, Inc.
            4633 World Parkway Circle, St. Louis, MO 63134, USA
            Tel: 877-577-2733
            Fax: 314-427-3324
            Email: service@genomesystems.com.
FEATURES             Location/Qualifiers
     source          1. .552
                     /organism="Arabidopsis thaliana"
                     /cultivar="Columbia Col-0"
                     /db_xref="taxon:3702"
                     /clone="701545727"
                     /clone_lib="A. thaliana, Columbia Col-0, rosette-2"
                     /tissue_type="rosette"
                     /dev_stage="4 - 7 weeks"
                     /note="Vector: pSPORT; Site_1: NotI; Site_2: SalI; cDNA
                     library was derived from untreated rosette tissue from
                     Arabidopsis  thaliana, Columbia Col-0, at 4 - 7 weeks.
                     Plants were grown in 1:1:1 peat  moss/vermiculite/perlite
                     soil at 22 deg. C +/- 3 deg. C under constant light,  and
                     watered with fertilizer. cDNA synthesis was initiated
                     using a NotI-oligo(dT)  primer. Double-stranded cDNA was
                     blunted, ligated to SalI adaptors, digested  with NotI,
                     size-selected, and cloned into the NotI and SalI sites of
                     the pSPORT vector."
BASE COUNT      125 a    134 c     92 g    178 t     23 others
ORIGIN


  Query Match             41.4%;  Score 504.4;  DB 9;  Length 552;
  Best Local Similarity   95.3%;  Pred. No. 3.9e-61;
  Matches  527;  Conservative    0;  Mismatches   24;  Indels    2;  Gaps    2;

Qy    598 gatgaagatgactcagatgatgacgaggaggaaggattctgaggatgaagaagaggagga 657
```

```
                    for salt-induced transcripts; 10-14 day seedlings; 4h
                    160mM NaCl stress"
                    /tissue_type="mixed leaf and root"
                    /dev_stage="10-14 day old"
                    /lab_host="Escherichia coli, TOP10F'"
                    /note="Organ: whole seedlings; Vector: pT-Adv (CLONTECH);
                    Site_1: EcoRI; Site_2: EcoRI; AD Arabidopsis thaliana
                    library enriched for salt-induced transcripts from
                    wildtype (Col-0 gl1) 10-14 day-old seedlings treated with
                    160 mM NaCl for 4 hours."
BASE COUNT      219 a    151 c    179 g    179 t
ORIGIN


  Query Match           50.8%;  Score 619.2;  DB 10;  Length 728;
  Best Local Similarity   96.1%;  Pred. No. 3.8e-77;
  Matches  688;  Conservative    0;  Mismatches   23;  Indels    5;  Gaps    5;

Qy      15 aaaaaaaacctaacaacctctcttctctcttcctcgttcaacaacaatggagttctgggg 74
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7 AAAAAAAACCTAACAACCTCTCTTCTCTCTTCCTCGTTCAACAACAATGGAGTTCTGGGG 66

Qy      75 agttgcggtgacaccaaaaaacgctactaaggtgactcctgaagaagacagccttgtcca 134
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67 AGTTGCGGTGACACCAAAAAACGCTACTAAGGTGACTCCTGAAGAAGACAGCCTTGTCCA 126

Qy     135 catttctcaggcttcacttgactgcacagtgaaatctggagaatctgtggttttgagtgt 194
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     127 CATTTCTCAGGCTTCACTTGACTGCACAGTGAAATCTGGAGAATCTGTGGTTTTGAGTGT 186

Qy     195 gactgttggtggggctaaacttgttattggaacactttcacaagacaagttccctcagat 254
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     187 GACTGTTGGTGGGGCTAAACTTGTTATTGGAACACTTTCACAAGACAAGTTCCCTCAGAT 246

Qy     255 tagctttgatttggtttttgataaagagtttgagctttcacacagcggtaccaaagcaaa 314
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     247 TAGCTTTGATTTGGTTTTTGATAAAGAGTTTGAGCTTTCACACAGCGGTACCAAAGCAAA 306

Qy     315 tgttcatttcattggctacaaatcccccaacatcgagcaggatgacttcactagttcgga 374
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     307 TGTTCATTTCATTGGCTACAAATCCCCCAACATCGAGCAGGATGACTTCACTAGTTCGGA 366

Qy     375 tgatgaggatgttcctgaagctgttcctgctcctgcccctactgctgttactgccaacgg 434
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     367 TGATGAGGATGTTCCTGAAGCTGTTCCTGCTCCTGCCCCTACTGCTGTTACTGCCAACGG 426

Qy     435 aaatgctggagcagctgttgtcaaggctgacacaaagccaaaggccaaacctgccgaagt 494
           ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db     427 AAATGCTGGAGCAGCTGTTGTCAAGGCTGACACATAGCCAAAGGCCAAACCTGCCGAAGT 486

Qy     495 gaagcctgcagaagagaagcctgaatcagacgaggaagatgagtctgatgatgaagatga 554
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     487 GAAGCCTGCAGAAGAGAAGCCTGAATCAGACGAGGAAGATGAGTCTGATGATGAAGATG- 545

Qy     555 agtctgaagaggatgatgactcttgagaaaggaatggatggttgatgaagatgactcaga 614
           |||||||||||||||||||| |||||||||||||||| ||||||||||||||||||||||
Db     546 AGTCTGAAGAGGATGATGACTC-TGAGAAAGGAATGGAT-GTTGATGAAGATGACTCAGA 603

Qy     615 tgatgacgaggaggaaggattctgaggatgaagaagaggaggagctcttaagaagcctga 674
           |||||||||||||| |||||||||| ||||||||||||||||||   |  |  |  |
Db     604 TGATGACGAGGAGG-AGGATTCTGACGATGAAGAAGAGGAGGAGACTTCTAAGAAGCCTG 662

Qy     675 gccaatcaacaagaagaggccaaatgaatctgtatccaaaacacccgtctctggaa 730
             | ||  ||| ||||| |||||||||||||||||| ||||||||||| || ||
Db     663 AGCCATTCACATGAAGAAGCCAAATGAATCTGTATCC-AAACACCCGTCTTTGAAA 717


RESULT  2
AV543093/c
LOCUS       AV543093                584 bp    mRNA    linear   EST 06-SEP-2000
DEFINITION  AV543093 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
            cDNA clone RZ196b09F 3', mRNA sequence.
ACCESSION   AV543093
VERSION     AV543093.1  GI:8704853
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 584)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     A large scale analysis of cDNA in Arabidopsis thaliana: Generation
            of 12,028 non-redundant expressed sequence tags from normalized and
            size-selected cDNA libraries
  JOURNAL   DNA Res. 7, 175-180 (2000)
  MEDLINE   20363093
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .584
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /clone="RZ196b09F"
                     /clone_lib="Arabidopsis thaliana roots Columbia"
                     /tissue_type="roots"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT      131 a    145 c    109 g    198 t      1 others
ORIGIN


  Query Match           45.6%;  Score 554.8;  DB 9;  Length 584;
  Best Local Similarity   99.1%;  Pred. No. 3.7e-68;
  Matches  578;  Conservative    0;  Mismatches    3;  Indels    2;  Gaps    2;

Qy     603 agatgactcagatgatgacgaggaggaaggattctgaggatgaagaagaggaggag-ctc 661
           ||||||||||||||||| |||||||| |||||||||||||||||||||||||||| ||| |
Db     584 AGATGACTCAGATGATGNCGAGGAGG-AGGATTCTGAGGATGAAGAAGAGGAGGAGACTC 526

Qy     662 ttaagaagcctgagccaatcaacaagaagaggccaaatgaatctgtatccaaaacacccg 721
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     525 CTAAGAAGCCTGAGCCAATCAACAAGAAGAGGCCAAATGAATCTGTATCCAAAACACCCG 466

Qy     722 tctctggaaagaaggcaaaaccagcagcagcaccagcttctactcctcagaagacagaag 781
           |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db     465 TCTCTGGAAAGAAGGCAAAACCAGCAGCAGCACCAGCTTCTACTCCTCAGAAGCCAGAAG 406

Qy     782 agaagaagaaaggaggacacaccgccacaccacacccagctaagaagggtggaaagtctc 841
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     405 AGAAGAAGAAAGGAGGACACACCGCCACACCACACCCAGCTAAGAAGGGTGGAAAGTCTC 346

Qy     842 ctgtgaatgctaaccagagccccaagtctggaggtcaatcatccggtggtaacaacaaca 901
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     345 CTGTGAATGCTAACCAGAGCCCCAAGTCTGGAGGTCAATCATCCGGTGGTAACAACAACA 286

Qy     902 agaagccattcaactcaggcaaacaatttggtggttccaacaacaagggttctaacaagg 961
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     285 AGAAGCCATTCAACTCAGGCAAACAATTTGGTGGTTCCAACAACAAGGGTTCTAACAAGG 226

Qy     962 gcaagggaaagggtagagcttaaggacgtggatcaaggagaggttttgggttttcgagta 1021
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     225 GCAAGGGAAAGGGTAGAGCTTAAGGACGTGGATCAAGGAGAGGTTTTGGGTTTTCGAGTA 166

Qy    1022 gatgatgaaaacacttggaagtgtggttttggatttttatcttatttttattagtataact 1081
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     165 GATGATGAAAACACTTGGAAGTGTGGTTTTGGATTTTTATCTTATTTTTATTAGTATAACT 106

Qy    1082 tgttatcggatgagctattttgagtatttgcaatttctacttttcctatgtaattcagtat 1141
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105 TGTTATCGGATGAGCTATTTTGAGTATTTGCAATTTCTACTTTCCTATGTAATTCAGTAT 46
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 18:43:44 ; Search time 3265.32 Seconds
(without alignments)
5034.512 Million cell updates/sec

Title: US-09-645-337-7
Perfect score: 1218
Sequence: 1 gtctttcgcttctaaaaaaa.....aaaaaaaaaagggcggccgc 1218

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	619.2	50.8	728	10	BE844690	BE844690 AD01A07T7
c	2	554.8	45.6	584	9	AV543093	AV543093 AV543093
c	3	534.8	43.9	559	9	AV543549	AV543549 AV543549
c	4	504.4	41.4	552	9	AI998449	AI998449 701545727
c	5	502	41.2	503	9	AV542413	AV542413 AV542413
c	6	479.8	39.4	483	9	AV559263	AV559263 AV559263
c	7	475.4	39.0	477	9	AV565954	AV565954 AV565954
c	8	458.4	37.6	585	9	AW004431	AW004431 701931423
c	9	426	35.0	426	9	AV539421	AV539421 AV539421
	10	412.4	33.9	422	10	BE523968	BE523968 M44A9STM
c	11	411.4	33.8	421	9	AV542939	AV542939 AV542939
	12	409.4	33.6	504	10	T45874	T45874 9137 Lambda
	13	407	33.4	444	9	AA597939	AA597939 29188 Lam
c	14	406.4	33.4	422	9	AV537871	AV537871 AV537871
c	15	404	33.2	407	9	AV541169	AV541169 AV541169
c	16	383	31.4	383	9	AV550412	AV550412 AV550412
	17	376.4	30.9	378	9	AV551556	AV551556 AV551556
	18	363.4	29.8	382	10	C99809	C99809 C99809 YAC
	19	354.6	29.1	457	9	AA713170	AA713170 32730 Lam
c	20	351	28.8	387	9	AV539161	AV539161 AV539161
	21	327	26.8	327	9	AV549937	AV549937 AV549937
	22	274.6	22.5	432	9	AA395217	AA395217 27000 Lam
c	23	267.6	22.0	282	9	AV539637	AV539637 AV539637
c	24	263	21.6	265	10	Z35338	Z35338 ATTS3777 AC
	25	249	20.4	249	10	BE525763	BE525763 M63018STM
	26	238.8	19.6	298	10	Z34946	Z34946 ATTS3623 St
	27	226	18.6	237	10	Z35337	Z35337 ATTS3776 AC
	28	221.2	18.2	333	10	Z24526	Z24526 ATTS968 Ral
	29	213	17.5	263	10	Z17912	Z17912 ATTS0396 AC
	30	194	15.9	288	10	T43469	T43469 6732 Lambda
	31	184	15.1	220	10	Z17565	Z17565 ATTS0139 AC
	32	147.4	12.1	534	9	AI995198	AI995198 701502609
	33	147.4	12.1	715	10	BF270691	BF270691 GA__Eb000
	34	142.2	11.7	543	10	BF279194	BF279194 GA__Eb003
	35	139.2	11.4	434	10	N37851	N37851 19078 Lambd
	36	139.2	11.4	875	10	BF275467	BF275467 GA__Eb002
	37	135.4	11.1	865	10	BG445639	BG445639 GA__Ea002
	38	125.2	10.3	390	9	AU068818	AU068818 AU068818
	39	125.2	10.3	517	10	BF278973	BF278973 GA__Eb003
	40	123.2	10.1	458	10	C72062	C72062 C72062 Rice
	41	122.4	10.0	787	9	AU082425	AU082425 AU082425
	42	121.2	10.0	397	9	AU067990	AU067990 AU067990
	43	117.8	9.7	692	10	BI543491	BI543491 S2S3A_H3
	44	114.8	9.4	742	10	BG526132	BG526132 57-3 Stev
	45	113.8	9.3	402	10	D15380	D15380 RICC0557A R

ALIGNMENTS

RESULT 1
BE844690
LOCUS BE844690 728 bp mRNA linear EST 17-MAY-2001
DEFINITION AD01A07T7 AD A. thaliana (Col-0 gl1) library enriched for salt-induced transcripts; 10-14 day seedlings; 4h 160mM NaCl stress Arabidopsis thaliana cDNA clone AD01A07 similar to (AF044914) putative histone deacetylase, mRNA sequence.
ACCESSION BE844690
VERSION BE844690.1 GI:10277068
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 728)
AUTHORS Gong,Z., Koiwa,H., Cushman,M.A., Ray,A., Bufford,D., Kore-Eda,S., Matsumoto,T.K., Zhu,J., Cushman,J.C., Bressan,R.A. and Hasegawa ,P.M.
TITLE Genes that are uniquely stress regulated in salt overly sensitive (sos) mutants
JOURNAL Plant Physiol. 126 (1), 363-375 (2001)
MEDLINE 21249177
COMMENT Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
Plate: AD01 row: A column: 7
Seq primer: T7
POLYA-No.
FEATURES Location/Qualifiers
source 1. .728
/organism="Arabidopsis thaliana"
/strain="ecotype Columbia gl1"
/db_xref="taxon:3702"
/clone="AD01A07"
/clone_lib="AD A. thaliana (Col-0 gl1) library enriched

```
Qy    451 gttgtcaaggctgacacaaagccaaaggccaaacctgccgaagtgaagcctgcagaagag 510
          |||| ||||| |||||   |    ||||   |    |  ||   | || | |  || |||
Db  21036 GTTGACAAGGATGACAATGACAATAAGGATGATGAGGAGGAGCAGGAGACAGATGAGGAG 20977

Qy    511 aagcctgaatcagacgaggaagatgagtctgatgatgaagatgaagtctgaagaggatga 570
          |   ||    |||||||| |||||    || ||||| || || |  ||| ||||| ||
Db  20976 GACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGA-TGACGAGGAGGA 20918

Qy    571 tgactcttgagaa-aggaatggatggttgatgaagatgactcagatgatgacgaggagga 629
          ||||      ||  | ||   |   |  || || ||||||   || |||||,|||||||
Db  20917 TGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGG- 20857

Qy    630 aggattctgaggatgaagaagaggagga 657
          | ||    |||||||| || ||||||||
Db  20858 ATGACGAGGAGGATGACGAGGAGGAGGA 20831


RESULT 15
US-08-757-669A-20/c
; Sequence 20, Application US/08757669A
; Patent No. 6183751
;  GENERAL INFORMATION:
;    APPLICANT:  Chang, Yuan
;    APPLICANT:  Bohenzky, Roy A.
;    APPLICANT:  Russo, James J.
;    APPLICANT:  Edelman, Isidore S.
;    APPLICANT:  Moore, Patrick S.
;    TITLE OF INVENTION:  UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
;    TITLE OF INVENTION:  SEQUENCES AND USES THEREOF
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/757,669A
;      FILING DATE:
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  45185-F
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 278-0400
;      TELEFAX:  (212) 391-0525
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  32207 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-757-669A-20


  Query Match             4.0%;  Score 48.8;  DB 4;  Length 32207;
  Best Local Similarity   61.5%;  Pred. No. 0.015;
  Matches  128;  Conservative    0;  Mismatches   77;  Indels    3;  Gaps    3;

Qy    451 gttgtcaaggctgacacaaagccaaaggccaaacctgccgaagtgaagcctgcagaagag 510
          |||| ||||| |||||   |    ||||   |    |  ||   | || | |  || |||
Db  21036 GTTGACAAGGATGACAATGACAATAAGGATGATGAGGAGGAGCAGGAGACAGATGAGGAG 20977

Qy    511 aagcctgaatcagacgaggaagatgagtctgatgatgaagatgaagtctgaagaggatga 570
          |   ||    |||||||| |||||    || ||||| || || |  ||| ||||| ||
Db  20976 GACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGA-TGACGAGGAGGA 20918

Qy    571 tgactcttgagaa-aggaatggatggttgatgaagatgactcagatgatgacgaggagga 629
          ||||      ||  | ||   |   |  || || ||||||   || |||||||||||||
Db  20917 TGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGG- 20857

Qy    630 aggattctgaggatgaagaagaggagga 657
          | ||    |||||||| || ||||||||
Db  20858 ATGACGAGGAGGATGACGAGGAGGAGGA 20831


Search completed: May  2, 2002, 20:16:07
Job time: 15154 sec
```

```
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA to mRNA
;     ORIGINAL SOURCE:
;       ORGANISM:  Mus musculus
;       STRAIN:  osteoblastic cell line MC3T3E1
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  284..2671
US-08-738-349-1


  Query Match             4.1%;  Score 49.6;  DB 2;  Length 3581;
  Best Local Similarity   65.2%;  Pred. No. 0.0041;
  Matches   73;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy   1097 tattttgagtatttgcaatttctactttcctatgtaattcagtatatgaatatttgctga 1156
          ||||||  | |  |   | |  | ||||  ||||   |||| |||| |||       |
Db   3421 TATTTTATGAACATTGTACTGTTGCTTTAATATGAGCTTCAATATAAGAAGCAACCTTTG 3480

Qy   1157 aatgagaaagaagactcgaattgcaaacaaaaaaaaaaaaaaaaaaaaaaaa 1208
          ||  | ||| |||| ||   ||  ||| ||||||||||||||||||||||||
Db   3481 AAATAAAAAAAAAGATTCTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3532


RESULT  13
US-08-728-323A-1
; Sequence 1, Application US/08728323A
; Patent No. 5948676
;  GENERAL INFORMATION:
;    APPLICANT:  Chang, Yuan
;    APPLICANT:  Bohenzky, Roy A.
;    APPLICANT:  Russo, James J.
;    APPLICANT:  Edelman, Isidore S.
;    APPLICANT:  Moore, Patrick S.
;    TITLE OF INVENTION:  Immediate Early Protein From Kaposi's
;    TITLE OF INVENTION:  Sarcoma-Associated Herpesvirus, DNA
;    TITLE OF INVENTION:  Encoding Same And Uses Thereof
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/728,323A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  0575/52268/JPW/MSC/SKS
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-278-0400
;      TELEFAX:  212-391-0525
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3489 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..3489
US-08-728-323A-1


  Query Match             4.0%;  Score 48.8;  DB 2;  Length 3489;
  Best Local Similarity   61.5%;  Pred. No. 0.0064;
  Matches  128;  Conservative    0;  Mismatches   77;  Indels    3;  Gaps    3;

Qy    451 gttgtcaaggctgacacaaagccaaaggccaaacctgccgaagtgaagcctgcagaagag 510
          |||| ||||| |||||    |    ||||    |    |  ||    | || | |  || |||
Db    961 GTTGACAAGGATGACAATGACAATAAGGATGATGAGGAGGAGCAGGAGACAGATGAGGAG 1020

Qy    511 aagcctgaatcagacgaggaagatgagtctgatgatgaagatgaagtctgaagaggatga 570
          |    ||    |||||||| |||||    || ||||| || || |  ||| ||||| ||
Db   1021 GACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGA-TGACGAGGAGGA 1079

Qy    571 tgactcttgagaa-aggaatggatggttgatgaagatgactcagatgatgacgaggagga 629
          ||||      ||  | ||   |   |  || || ||||||   || ||||||||||||||
Db   1080 TGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGG- 1138

Qy    630 aggattctgaggatgaagaagaggaggga 657
          | ||    |||||||| || ||||||||
Db   1139 ATGACGAGGAGGATGACGAGGAGGAGGA 1166


RESULT  14
US-08-770-379-20/c
; Sequence 20, Application US/08770379
; Patent No. 5849564
;  GENERAL INFORMATION:
;    APPLICANT:  Chang, Yuan
;    APPLICANT:  Bohenzky, Roy A.
;    APPLICANT:  Russo, James J.
;    APPLICANT:  Edelman, Isidore S.
;    APPLICANT:  Moore, Patrick S.
;    TITLE OF INVENTION:  POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED
;    TITLE OF INVENTION:  HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/770,379
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  52342
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 278-0400
;      TELEFAX:  (212) 391-0525
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  32207 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-770-379-20


  Query Match             4.0%;  Score 48.8;  DB 2;  Length 32207;
  Best Local Similarity   61.5%;  Pred. No. 0.015;
  Matches  128;  Conservative    0;  Mismatches   77;  Indels    3;  Gaps    3;
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/867,106
;      FILING DATE:  19920625
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PJ 7187
;      APPLICATION NUMBER:  PCT/AU90/00530
;      FILING DATE:  02-NOV-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Feeney, Joanne Longo
;      REGISTRATION NUMBER:  35,134
;      REFERENCE/DOCKET NUMBER:  RICE-0002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5852 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  2378..5038
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  2378..5038
US-07-867-106-2


  Query Match             4.1%;  Score 50;  DB 1;  Length 5852;
  Best Local Similarity   57.8%;  Pred. No. 0.0039;
  Matches   89;  Conservative    0;  Mismatches   65;  Indels    0;  Gaps    0;

Qy   1055 tttttatcttattttattagtataacttgttatcggatgagctattttgagtatttgcaa 1114
          || |||| |   ||| ||  ||||  || ||    | |    | |||        || | |
Db   5758 TTGTTATTTGTATTTTTTTATATATGTTATTTGTTGTTTGTTTTTTTACTTATATTTCTA 5699

Qy   1115 tttctactttcctatgtaattcagtatatgaatatttgctgaaatgagaaagaagactcg 1174
          ||| |  |||   |  | | | | || ||    | |   | |||  | ||| || |
Db   5698 TTTTTTATTTATAAAATTAATAATTAAATTTTAAATAAATAAAAAAAAAAAAAAAAAAAA 5639

Qy   1175 aattgcaaacaaaaaaaaaaaaaaaaaaaaaaa 1208
          ||||  ||  ||||||||||||||||||||||||
Db   5638 AATTTAAATTAAAAAAAAAAAAAAAAAAAAAAAA 5605


RESULT  11
US-09-253-691-3/c
; Sequence 3, Application US/09253691
; Patent No. 6124100
; GENERAL INFORMATION:
;  APPLICANT: Dong Kyu JIN
;  TITLE OF INVENTION: Diagnostic Method and Kit for Neuropsychiatric Diseases
;  TITLE OF INVENTION: Using Trinucleotide Repeats Sequence
;  FILE REFERENCE: 1942/36
;  CURRENT APPLICATION NUMBER: US/09/253,691
;  CURRENT FILING DATE: 1999-02-22
;  EARLIER APPLICATION NUMBER: KR 98-6,278
;  EARLIER FILING DATE: 1996-02-26
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: WordPerfect 6.1/Windows
; SEQ ID NO 3
;   LENGTH: 397
;   TYPE: DNA
;   ORGANISM: human
US-09-253-691-3


  Query Match             4.1%;  Score 49.8;  DB 3;  Length 397;
  Best Local Similarity   49.8%;  Pred. No. 0.0016;
  Matches  126;  Conservative    0;  Mismatches  127;  Indels    0;  Gaps    0;

Qy    365 ctagttcggatgatgaggatgttcctgaagctgttcctgctcctgcccctactgctgtta 424
          ||  | | | || ||  | || | |||  |||| | ||||| ||||  || ||||||
Db    322 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 263

Qy    425 ctgccaacggaaatgctggagcagctgttgtcaaggctgacacaaagccaaaggccaaac 484
          ||||    |    |||||  || |||| ||     ||||   |     |    ||
Db    262 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 203

Qy    485 ctgccgaagtgaagcctgcagaagagaagcctgaatcagacgaggaagatgagtctgatg 544
          |||| |  |       |||| |  |       |||   | |  |  |  | ||   ||| ||
Db    202 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 143

Qy    545 atgaagatgaagtctgaagaggatgatgactcttgagaaaggaatggatggttgatgaag 604
           ||  | ||  |  |     |    |||  | |  ||||  | | ||  || | ||
Db    142 CTGCTGTTGCTGCTTTTGCTGCTGTCTGAAACATTCAAAAGTGAAGTATATTTAAAAAAC 83

Qy    605 atgactcagatga 617
          |  ||| | | ||
Db     82 AAAACTTAAAAGA 70


RESULT  12
US-08-738-349-1
; Sequence 1, Application US/08738349
; Patent No. 5869638
;  GENERAL INFORMATION:
;    APPLICANT:  Takeshita, Sunao
;    APPLICANT:  Okazaki, Makoto
;    APPLICANT:  Kawai, Shinji
;    APPLICANT:  Tsujimura, Atsushi
;    APPLICANT:  Amann, Egon
;    TITLE OF INVENTION:  Bone-Related Cadherin-Like Protein and
;    TITLE OF INVENTION:  Process for Its Production
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;      ADDRESSEE:  Dunner
;      STREET:  1300 I Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20005-3315
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/738,349
;      FILING DATE:  25-OCT-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/364,439
;      FILING DATE:
;      APPLICATION NUMBER:  US 08/112,061
;      FILING DATE:  26-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barker, M. P.
;      REGISTRATION NUMBER:  32,013
;      REFERENCE/DOCKET NUMBER:  02481.1323-00000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-408-4000
;      TELEFAX:  202-408-4400
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3581 base pairs
;      TYPE:  nucleic acid
```

```
RESULT   8
US-09-590-113-2
; Sequence 2, Application US/09590113
; Patent No. 6306389
; GENERAL INFORMATION:
;  APPLICANT: Tureci, Ozlem
;  APPLICANT:  Sakin, Ugur
;  APPLICANT:  Pfreundschuh, Michael
;  TITLE OF INVENTION: Isolated Nucleic Acid Molecules Which Encode SCP Proteins, And Us
;  TITLE OF INVENTION:  Thereof
;  FILE REFERENCE: LUD 5525
;  CURRENT APPLICATION NUMBER: US/09/590,113
;  CURRENT FILING DATE:  2000-06-08
;  PRIOR APPLICATION NUMBER: US 09/177,325
;  PRIOR FILING DATE: 1998-10-22
;  NUMBER OF SEQ ID NOS: 15
; SEQ ID NO 2
;   LENGTH: 1276
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
US-09-590-113-2


  Query Match             4.7%;  Score 57.8;  DB 4;  Length 1276;
  Best Local Similarity   51.3%;  Pred. No. 2.5e-05;
  Matches  134;  Conservative    0;  Mismatches  127;  Indels    0;  Gaps    0;

Qy     373 gatgatgaggatgttcctgaagctgttcctgctcctgcccctactgctgttactgccaac 432
           ||||| |  || |   | ||||| |   | ||  | ||     | || |   | ||  |
Db     645 gatgaagcagaagaagcagaagccgcagcagcagcagcggaagccgcagcagcagcagaa 704

Qy     433 ggaaatgctggagcagctgttgtcaaggctgacacaaagccaaaggccaaacctgccgaa 492
           |     || | ||||||| |  || |  |  |    | |   | |||   ||      ||
Db     705 gccgcagcagcagcagcagaagtaatagtagtagaagacgaagaggaggaagagaaggag 764

Qy     493 gtgaagcctgcagaagagaagcctgaatcagacgaggaagatgagtctgatgatgaagat 552
           | | ||   |  |||| ||  ||| ||| |  ||||| ||    |  |  |||||
Db     765 gaggaggaggagaaagaagaggaggaagaagaaggagaagaagaaggaggaggagaagaa 824

Qy     553 gaagtctgaagaggatgatgactcttgagaaaggaatggatggttgatgaaqatgactca 612
           | ||    | ||||| || ||       || | |||      |  || ||||| ||   |
Db     825 ggagaagaaggaggaggaggaggagaagaaggagaagaaacagaagaagaggaagaggaagaa 884

Qy     613 gatgatgacgaggaggaaagga 633
           || || || || |||||||
Db     885 gaagaagaggaagaggaagaa 905


RESULT   9
US-08-766-738-2
; Sequence 2, Application US/08766738
; Patent No. 5916749
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Goli, Surya K.
;    TITLE OF INVENTION:  NOVEL HUMAN PHOSPHATASE INHIBITOR PROTEIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/766,738
;      FILING DATE:  Herewith
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0177 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  966 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  Consensus
;      CLONE:  1813361
US-08-766-738-2


  Query Match             4.3%;  Score 52.8;  DB 2;  Length 966;
  Best Local Similarity   49.4%;  Pred. No. 0.0004;
  Matches  162;  Conservative    0;  Mismatches  164;  Indels    2;  Gaps    1;

Qy     374 atgatgaggatgttcctgaagctgttcctgctcctgcccctactgctgttactgccaacg 433
           ||||    || |  |  |||||   |    |  ||||      | || |  ||   | |
Db     533 ATGACCGAGAGGACCAGGAAGCACCTGACTCAGATGCCGAGGTGGATGGTGTTGATNAAG 592

Qy     434 gaaatgctggagcagctgttgtcaaggctgacacaaagccaaaggccaaacctgccgaag 493
             | |  |  | ||  |  |   | |  ||     ||    | |   |   ||  ||||
Db     593 ANGAGGAGGACGGAGAAGGAGAAGATGAGGAAGACGAGGACGATGAGGATGGTGAAGAAG 652

Qy     494 tgaagcctgcagaagagaagcctgaatcagacgaggaagatgagtctgatgatgaagatg 553
           | ||  ||  |||||  |   ||||   || || | ||| | |  ||| || || ||||
Db     653 AGGAGTTTGATGAAGAAGATGATGAAGATGAAGATGTAGAAGGGGATGAGGACGACGATG 712

Qy     554 aagtctgaagaggatgatgactcttgagaaaggaatggatggttgatgaaqatgactcag 613
           |||||  |    |   |  |  ||  ||  ||       | ||| || || ||    |
Db     713 AAGTCAGTGAGGAGGAAGAAGAATTTGGACTTGATGAAGAAGATGAAGATGAGGATGAGG 772

Qy     614 atgatgacgaggaggaaggattctgaggatgaagaagaggaggagctcttaagaagcctg 673
           |||| || ||||| |||||  |  ||      | |||||| |||      ||| |
Db     773 ATGAAGAGGAGGAAGAAGG--TGGGAAAGGTGAAAAGAGGAAGAGAGAAACAGATGATGA 830

Qy     674 agccaatcaacaagaagaggccaaatga 701
           ||   |  |  | ||||  ||| ||||
Db     831 AGGAGAAGATGATTAAGACCCCAGATGA 858


RESULT  10
US-07-867-106-2/c
; Sequence 2, Application US/07867106
; Patent No. 5389526
;  GENERAL INFORMATION:
;    APPLICANT:  Slade, Martin B
;    APPLICANT:  Chang, Andy C M
;    APPLICANT:  Williams, Keith L
;    TITLE OF INVENTION:  Improved Plasmid Vectors for Cellular
;    TITLE OF INVENTION:  Slime Moulds of the Genus Dictyostelium
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
;      STREET:  One Liberty Place 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
```

```
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14


  Query Match             7.0%;  Score 85.4;  DB 1;  Length 7218;
  Best Local Similarity   6.2%;  Pred. No. 6.8e-12;
  Matches   26;  Conservative  246;  Mismatches  147;  Indels    0;  Gaps    0;

Qy     422 ttactgccaacggaaatgctggagcagctgttgtcaaggctgacacaaagccaaaggcca 481
           ||  | | :: ::::: :  :::: ::  :  :  :::: :: : ::::  ::::: :
Db    1441 TTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1382

Qy     482 aacctgccgaagtgaagcctgcagaagagaagcctgaatcagacgaggaagatgagtctg 541
           ::   :  :::: ::::   : ::::::::::   :::  ::: :::::::: :::   :
Db    1381 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1322

Qy     542 atgatgaagatgaagtctgaagaggatgatgactcttgagaaaggaatggatggttgatg 601
           : :: ::::: ::::   :::::::: :: ::     :::::::::: ::: ::  :: :
Db    1321 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1262

Qy     602 aagatgactcagatgatgacgaggaggaaggattctgaggatgaagaagaggaggagctc 661
           :::: ::   ::: :: :: :::::::::::::    ::::: ::::::::::::::::
Db    1261 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1202

Qy     662 ttaagaagcctgagccaatcaacaagaagaggccaaatgaatctgtatccaaaacacccg 721
             ::::::   :::  ::  :: :::::::::  ::: :::   : :   :::: :   :
Db    1201 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1142

Qy     722 tctctggaaagaaggcaaaaccagcagcagcaccagcttctactcctcagaagacagaag 781
           :::::::::: ::::  :: :: :: :  ::      :      :::::: :::::
Db    1141 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1082

Qy     782 agaagaagaaaggaggacacaccgccacaccacacccagctaagaagggtggaaagtct 840
           ::::::::::::::  | ||    | | | ||   |||| |||   ||    || |||
Db    1081 RRRRRRRRRRRRRRRATCGCAAGCTCCCTCGACCTGCAGCCAAGCTCGGAATTAATTCT 1023


RESULT   6
US-09-177-325-2
; Sequence 2, Application US/09177325B
; Patent No. 6214983
; GENERAL INFORMATION:
;  APPLICANT:      Tureci, Ozlem
;  APPLICANT:  Sakin, Ugur
;  APPLICANT:  Pfreundschuh, Michael
;  TITLE OF INVENTION:      Isolated Nucleic Acid Molecules Which Encode SCP Proteins, A
;  TITLE OF INVENTION:  Thereof
;  FILE REFERENCE:      LUD 5525
;  CURRENT APPLICATION NUMBER: US/09/177,325B
;  CURRENT FILING DATE:  1998-10-22
;  NUMBER OF SEQ ID NOS:      15
; SEQ ID NO 2
;   LENGTH: 1276
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
US-09-177-325-2


  Query Match             4.7%;  Score 57.8;  DB 4;  Length 1276;
  Best Local Similarity   51.3%;  Pred. No. 2.5e-05;
  Matches  134;  Conservative    0;  Mismatches  127;  Indels    0;  Gaps    0;

Qy     373 gatgatgaggatgttcctgaagctgttcctgctcctgcccctactgctgttactgccaac 432
           ||||| |  || |   | ||||| |   | ||  | ||     | || |   | ||  |
Db     645 gatgaagcagaagaagcagaagccgcagcagcagcagcggaagccgcagcagcagcagaa 704

Qy     433 ggaaatgctggagcagctgttgtcaaggctgacacaaagccaaaggccaaacctgccgaa 492
           |      || | |||||| |  || |  | |    | |   | |||   ||      ||
Db     705 gccgcagcagcagcagcagaagtaatagtagtagaagacgaagaggaggaagagaaggag 764

Qy     493 gtgaagcctgcagaagagaagcctgaatcagacgaggaagatgagtctgatgatgaagat 552
           | | ||   |   ||||  ||   |||  ||| | ||||| ||    |  |  |||||
Db     765 gaggaggaggagaaagaagaggaggaagaagaaggagaagaagaaggaggaggagaagaa 824

Qy     553 gaagtctgaagaggatgatgactcttgagaaaggaatggatggttgatgaagatgactca 612
           | ||    | ||||| || ||       || | |||      |  || |||||  ||   |
Db     825 ggagaagaaggaggaggaggaggagaagaggagaagaaacagaagaagaggaagaggaaga 884

Qy     613 gatgatgacgaggaggaagga 633
           || || || || |||||||
Db     885 gaagaagaggaagaggaagaa 905


RESULT   7
US-09-411-812A-2
; Sequence 2, Application US/09411812A
; Patent No. 6261778
; GENERAL INFORMATION:
;  APPLICANT:      Tureci, Ozlem
;  APPLICANT:  Sahin, Ugur
;  APPLICANT:  Pfreundschuh, Michael
;  TITLE OF INVENTION:   Isolated Nucleic Acid Molecules Which Encode SCP Proteins, A
;  TITLE OF INVENTION:  Thereof
;  FILE REFERENCE:      LUD 5525.1 CIP
;  CURRENT APPLICATION NUMBER: US/09/411,812A
;  CURRENT FILING DATE:  1999-10-01
;  PRIOR APPLICATION NUMBER:      US 09/177,325
;  PRIOR FILING DATE:      1998-10-22
;  NUMBER OF SEQ ID NOS:      14
; SEQ ID NO 2
;   LENGTH: 1276
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-411-812A-2


  Query Match             4.7%;  Score 57.8;  DB 4;  Length 1276;
  Best Local Similarity   51.3%;  Pred. No. 2.5e-05;
  Matches  134;  Conservative    0;  Mismatches  127;  Indels    0;  Gaps    0;

Qy     373 gatgatgaggatgttcctgaagctgttcctgctcctgcccctactgctgttactgccaac 432
           ||||| |  || |   | ||||| |   | ||  | ||     | || |   | ||  |
Db     645 gatgaagcagaagaagcagaagccgcagcagcagcagcggaagccgcagcagcagcagaa 704

Qy     433 ggaaatgctggagcagctgttgtcaaggctgacacaaagccaaaggccaaacctgccgaa 492
           |      || | |||||| |  || |  | |    | |   | |||   ||      ||
Db     705 gccgcagcagcagcagcagaagtaatagtagtagaagacgaagaggaggaagagaaggag 764

Qy     493 gtgaagcctgcagaagagaagcctgaatcagacgaggaagatgagtctgatgatgaagat 552
           | | ||   |   ||||  ||   |||  ||| | ||||| ||    |  |  |||||
Db     765 gaggaggaggagaaagaagaggaggaagaagaaggagaagaagaaggaggaggagaagaa 824

Qy     553 gaagtctgaagaggatgatgactcttgagaaaggaatggatggttgatgaagatgactca 612
           | ||    | ||||| || ||       || | |||      |  || |||||  ||   |
Db     825 ggagaagaaggaggaggaggaggagaagaggagaagaaacagaagaagaggaagaggaaga 884

Qy     613 gatgatgacgaggaggaagga 633
           || || || || |||||||
Db     885 gaagaagaggaagaggaagaa 905
```

```
Qy    313 aatgttcatttcattggctacaaatcccccaacatcgagcaggatgacttcactagttcg 372
          | |||   ||||  ||| |||||         |        |||  ||    |    |
Db    333 agtgtatttttctctggttacaaggttgagcagccaattgagggagatgaaatggatctt 392

Qy    373 gatgatgaggatgttcctgaagctgttcctgctcctg-----------ccctactgct 420
          |||  ||||||||     || |   |      ||| |                |  |||
Db    393 gattctgaggatgaagaggaggagctaaacattccagtaatcaaggaaaatggcaaagct 452

Qy    421 gttactgccaacggaaatgctggagcagctgttgtcaaggctgacacaaagccaaaggcc 480
          | |       | |   |          ||     |     ||||  |||    ||||  |
Db    453 gatgggaaggaggagcagaaaaatcaagagaaggcagtagctgctacagcttcaaaatca 512

Qy    481 aaacctgccgaagtgaagcctgcagaagagaagcctgaatcagacgaggaagatgagtct 540
          |  | || |   |  |||       | ||  |  ||||       || || |||||||||
Db    513 agtcttggccttgaaaagaaaagcaaggatgactctgatgattctgatgaggatgagtct 572

Qy    541 gatgatgaagatgaagtctgaagaggatgatgactcttgagaaaggaatggat------- 593
          ||||||   |||||     || ||   |||||| |||   ||| |  | ||||
Db    573 gatgattctgatga----ggatgattctgatgattctgatgaaggcgagggattatctcc 628

Qy    594 ---ggttgatgaagatgactcagatgatgacgaggaggaaggattctgaggatgaagaag 650
              |  |  || ||||| |||  |||||| || ||     |      |       |
Db    629 tgacgaaggcgatgatgattcaagtgatgaggatgataccagtgatgatgacgaggaaga 688

Qy    651 aggaggagctcttaagaagcctgagccaatcaacaagaagaggccaaatgaatctgtatc 710
          |     | ||| ||| ||||| ||| ||     |||||||||      ||||  ||
Db    689 aaccccaactcctaaaaagccagaggca---ggcaagaagagaggtgctgaaaatgctct 745

Qy    711 caaaacacccgtctctggaaagaaggcaaaaccagcagcagcaccagcttctactcctca 770
           ||||| ||  | ||||  ||||| |||||    ||  || | ||||         | ||
Db    746 gaaaacgcctctttctgataagaaagcaaaggttgccacaccgccag---------ccca 796

Qy    771 gaagacagaagagaagaagaaaggaggacacaccgccacaccacacccagctaa 824
          ||| ||||  |  ||||||   |    ||    || || ||||||||||| ||
Db    797 gaaaacaggtggcaagaagggtgccacccatgtggcaactccacacccagcaaa 850


RESULT   4
US-09-282-305-11
; Sequence 11, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
;  APPLICANT:  Baldwin, Donald A.
;  APPLICANT:  Briggs, Steven P.
;  APPLICANT:  Crane, Virginia C.
;  TITLE OF INVENTION:  Maize Histone Deacetylases And Their Uses
;  FILE REFERENCE:  5718-44,
;  CURRENT APPLICATION NUMBER: US/09/282,305
;  CURRENT FILING DATE:  1999-03-31
;  PRIOR APPLICATION NUMBER:  60/080,563
;  PRIOR FILING DATE:  1998-04-03
;  NUMBER OF SEQ ID NOS:  18
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 11
;   LENGTH: 1283
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (79)..(996)
US-09-282-305-11


  Query Match             9.1%;  Score 111.2;  DB 4;  Length 1283;
  Best Local Similarity   50.7%;  Pred. No. 1.4e-18;
  Matches  359;  Conservative    0;  Mismatches  328;  Indels   21;  Gaps    3;

Qy     61 atggagttctggggagttgcggtgacaccaaaaaacgctactaaggtgactcctgaagaa 120
          ||||||||||||||  | | ||| |  ||      | ||  ||||     |||| | |
Db     79 atggagttctggggtctcgaggtcaagcctggttccactgttaagtgtgagcctggatat 138

Qy    121 gacagccttgtccacatttctcaggcttcacttgactgcacagtgaaatctggagaatct 180
          | |    |  | ||| |||| |||||| | ||||   |   |  |||     | ||   |

Db    139 ggctttgtgctgcacctttcccaggctgctcttg---gggaatcgaagaagagtgataat 195

Qy    181 gtggttttgagtgtgactgttggtggggctaaacttgttattggaacactttcacaagac 240
          |   |  ||  ||| |   ||| ||     |||||||  |||||||| || ||    |||
Db    196 gccttgatgtatgtcaaaattgatgatcagaaacttgccattggaaccctctctgttgac 255

Qy    241 aagttccctcagattagctttgatttggttttgataaagagtttgagctttcacacagc 300
          |||  ||| || ||    |||||| || |||| ||||||||||||||||| || ||||
Db    256 aagaacccacacatacaatttgatctgattttcgataaagagtttgagctgtcgcacaca 315

Qy    301 ggtaccaaagcaaatgttcatttcattggctacaaatcccccaacatcgagcaggatgac 360
          |  |  | |  |||   |||| |||||||||        |       |||| ||
Db    316 tcaaaaactaccagcgttttcttcactggctacaaggttgagcagccattcgaggaagat 375

Qy    361 ttcactagttcggatgatgaggatgttcctgaagctgttcctgctcctgcccctactgct 420
          |    |  |||  ||| ||||     || |   |   || ||         | |
Db    376 gaaatggatcttgattctgaagatgaagacgaggagctgaatgttc---------cagta 426

Qy    421 gttactgccaacggaaatgctggagcagctgttgtcaaggctgacacaaagccaaaggcc 480
          || |  |  || || || ||||  |           ||   | |   |||| ||   ||
Db    427 gtcaaggaaaatggcaaagctgatgagaagaaacagaaaagtcaagaaaaggcagttgct 486

Qy    481 aaacctgccgaagtgaagcctgcagaagagaagcctgaatcagac---------gaggaa 531
          ||||  |  ||   |  || |      |||||    |    |||         || ||
Db    487 gcaccttcaaaatcaagtccggattccaagaagagcaaggatgacgacgattctgatgag 546

Qy    532 gatgagtctgatgatgaagatgaagtctgaagaggatgatgactcttgagaaaggaatgg 591
          || ||| ||||||||   ||||| |     |  |     |       | |   |
Db    547 gacgagactgatgattctgatgaggatgagacggacgattctgatgagggtttgtctcct 606

Qy    592 atggttgatgaagatgactcagatgatgacgaggaggaaggattctgaggatgaagaaga 651
          |  |  || ||||| |||  |||||| || ||     |   |   |       |
Db    607 gaagaaggcgatgatgattcaagtgatgaagatgataccagtgacgatgaggaggaagac 666

Qy    652 ggaggagctcttaagaagcctgagccaatcaacaagaagaggccaaatgaatctgtatcc 711
          | ||| ||||||||||||||  |  ||| ||||     |  |    || ||
Db    667 actccaactcctaagaagcctgaggtaggcaagaagagagctgctgaaagttccgtgctg 726

Qy    712 aaaacacccgtctctggaaagaaggcaaaaccagcagcagcaccagct 759
          ||||| ||  | ||||  ||||| |||||    ||  || |  || ||
Db    727 aaaactcctctatctgataagaaagcaaaggttgccacaccgtcatct 774


RESULT   5
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
```

```
Qy    357 -------tgacttcactagttcggatgatgaggatgttcctgaagctgttcctgctcc-- 407
Db    385 ctgattctggtgaagatagttcagatgaagaggttgaacccgatctaattccaatgcaga 444

Qy    408 ----tgcccctactgctgttactgccaacggaaatgctggagcagctgttgtcaaggctg 463
Db    445 ataatgaaattaaaatttctactgcaaaggttcccgtgaaggttggtatacaaaatgctg 504

Qy    464 acacaaagccaaaggccaaacctgccgaagtgaagcctgcagaagagaagcctgaatcag 523
Db    505 atgaagatgaaacttctagtggtgatgatgatgatttcactgatagtgatagtgaaatgt 564

Qy    524 acgaggaagatgagtctgatgatgaagatgaagtctgaagag 565
Db    565 ctgaggaagatgagtccagtgatgaagatgaagtgtcaagcg 606
```

```
RESULT   2
US-09-282-305-13
; Sequence 13, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
;  APPLICANT:  Baldwin, Donald A.
;  APPLICANT:  Briggs, Steven P.
;  APPLICANT:  Crane, Virginia C.
;  TITLE OF INVENTION:  Maize Histone Deacetylases And Their Uses
;  FILE REFERENCE:  5718-44,
;  CURRENT APPLICATION NUMBER: US/09/282,305
;  CURRENT FILING DATE:  1999-03-31
;  PRIOR APPLICATION NUMBER:  60/080,563
;  PRIOR FILING DATE:  1998-04-03
;  NUMBER OF SEQ ID NOS:  18
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 13
;   LENGTH: 1191
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (63)..(971)
US-09-282-305-13
```

```
  Query Match             9.2%;  Score 111.6;  DB 4;  Length 1191;
  Best Local Similarity   51.4%;  Pred. No. 1.1e-18;
  Matches  362;  Conservative    0;  Mismatches  329;  Indels   13;  Gaps    4;

Qy     61 atggagttctggggagttgcggtgacaccaaaaaacgctactaaggtgactcctgaagaa 120
Db     63 atggagttctggggtctcgaggtcaagcctggttccactgttaagtgtgagcctggatat 122

Qy    121 gacagccttgtccacatttctcaggcttcacttgactgcacagtgaaatctggagaatct 180
Db    123 ggctttgtgctgcacctttcccaggctgctcttg---gggaatcgaagaagagtgataat 179

Qy    181 gtggttttgagtgtgactgttggtggggctaaacttgttattggaacactttcacaagac 240
Db    180 gccttgatgtatgtcaaaattgatgatcagaaacttgccattggaaccctctctgttgac 239

Qy    241 aagttccctcagattagctttgatttggtttttgataaagagtttgagctttcacacagc 300
Db    240 aagaacccacacattcaatttgatctgattttcgataaagagtttgagctgtcgcacaca 299

Qy    301 ggtaccaaagcaaatgttcatttcattggctacaaatcccccaacatcgagcaggatgac 360
Db    300 tcaaaaactaccagcgtcttcttcactggctacaaggttgaacagccattcgaggaagat 359

Qy    361 ttcactagttcggatgatgaggatgttcctgaagctgttcctgctcctgcccctactg-- 418
Db    360 gaaatggatcttgattctgaagatgaagacgaggagctgaatgttccagtagtcaaggaa 419

Qy    419 -ctgttactgccaacggaaatgctggagcagctgttgtcaaggctgacacaaagccaaag 477
```

```
Db    420 aatggcaaagctgatgggaagaaacagaaaagtcaagaaaaggcagttgctgcaccttca 479

Qy    478 gccaaacctgccgaagtgaagcctgcagaagagaagcctgaatcagacgaggaagatgag 537
Db    480 aaatcaagtccggattccaagaagagcaaggatgacgatgattctgatgaggacgagact 539

Qy    538 tctgatgatgaagatga--agtctgaagaggatgatgactcttgagaaaggaatggatgg 595
Db    540 gatgattctgatgaggatgagacggacgattctgatgagggtttatcttctgaagaa--- 596

Qy    596 ttgatgaagatgactcagatgatgacgaggaggaaggattctgaggatgaagaagaggag 655
Db    597 --ggcgatgatgattcaagtgatgaagatgataccagtgacgatgaggaggaagacactc 654

Qy    656 gagctcttaagaagcctgagccaatcaacaagaagaggccaaatgaatctgtatccaaaa 715
Db    655 caactcctaagaagcctgaggtaggcaagaagagagctgctgaaagttccgtgctgaaaa 714

Qy    716 cacccgtctctggaaagaaggcaaaaccagcagcagcaccagct 759
Db    715 ctcctctatctgataagaaagcaaaggttgccacaccgtcatct 758
```

```
RESULT   3
US-09-282-305-15
; Sequence 15, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
;  APPLICANT:  Baldwin, Donald A.
;  APPLICANT:  Briggs, Steven P.
;  APPLICANT:  Crane, Virginia C.
;  TITLE OF INVENTION:  Maize Histone Deacetylases And Their Uses
;  FILE REFERENCE:  5718-44,
;  CURRENT APPLICATION NUMBER: US/09/282,305
;  CURRENT FILING DATE:  1999-03-31
;  PRIOR APPLICATION NUMBER:  60/080,563
;  PRIOR FILING DATE:  1998-04-03
;  NUMBER OF SEQ ID NOS:  18
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 15
;   LENGTH: 1245
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (84)..(1019)
US-09-282-305-15
```

```
  Query Match             9.1%;  Score 111.2;  DB 4;  Length 1245;
  Best Local Similarity   51.6%;  Pred. No. 1.4e-18;
  Matches  430;  Conservative    0;  Mismatches  363;  Indels   41;  Gaps    6;

Qy     13 taaaaaaaaacctaacaacctctcttctctcttcctcgttcaacaacaatggagttctgg 72
Db     36 ttaaaccctagttcccattccatcttcgtttccgctccgccgccgtcgatggagttctgg 95

Qy     73 ggagttgcggtgacaccaaaaaacgctactaaggtgactcctgaagaagacagccttgtc 132
Db     96 ggtctcgaggtcaaacctggatccactgtcaagtgtgagcctggacatggctttatcctg 155

Qy    133 cacatttctcaggcttcacttgactgcacagtgaaatctggagaatctgtggttttgagt 192
Db    156 cacgtttcccaggctgcccttggggaatcaaagaaaagt---gacagtgccttaatgtat 212

Qy    193 gtgactgttggtggggctaaacttgttattggaacactttcacaagacaagttccctcag 252
Db    213 gtcaaagttgatgacaagaagcttgccattggaacgctctctatcgacaaatacccacag 272

Qy    253 attagctttgatttggtttttgataaagagtttgagctttcacacagcggtaccaaagca 312
Db    273 atacaattcgatttggtttcaataaagagtttgagctgtcacacacatcgaaaactacc 332
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 20:15:18 ; Search time 100.71 Seconds
(without alignments)
2970.724 Million cell updates/sec

Title: US-09-645-337-7
Perfect score: 1218
Sequence: 1 gtctttcgcttctaaaaaaa.....aaaaaaaaaagggcggccgc 1218

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	112	9.2	1307	4	US-09-282-305-17	Sequence 17, Appl
	2	111.6	9.2	1191	4	US-09-282-305-13	Sequence 13, Appl
	3	111.2	9.1	1245	4	US-09-282-305-15	Sequence 15, Appl
	4	111.2	9.1	1283	4	US-09-282-305-11	Sequence 11, Appl
c	5	85.4	7.0	7218	1	US-08-232-463-14	Sequence 14, Appl
	6	57.8	4.7	1276	4	US-09-177-325-2	Sequence 2, Appli
	7	57.8	4.7	1276	4	US-09-411-812A-2	Sequence 2, Appli
	8	57.8	4.7	1276	4	US-09-590-113-2	Sequence 2, Appli
	9	52.8	4.3	966	2	US-08-766-738-2	Sequence 2, Appli
c	10	50	4.1	5852	1	US-07-867-106-2	Sequence 2, Appli
c	11	49.8	4.1	397	3	US-09-253-691-3	Sequence 3, Appli
	12	49.6	4.1	3581	2	US-08-738-349-1	Sequence 1, Appli
	13	48.8	4.0	3489	2	US-08-728-323A-1	Sequence 1, Appli
c	14	48.8	4.0	32207	2	US-08-770-379-20	Sequence 20, Appl
c	15	48.8	4.0	32207	4	US-08-757-669A-20	Sequence 20, Appl
c	16	48.8	4.0	32207	4	US-09-230-371A-20	Sequence 20, Appl
	17	48.6	4.0	1223	3	US-09-154-874-4	Sequence 4, Appli
	18	47.6	3.9	2625	4	US-09-245-041-18	Sequence 18, Appl
c	19	47	3.9	509	4	US-09-030-607-202	Sequence 202, App
c	20	47	3.9	509	4	US-09-439-313-202	Sequence 202, App
c	21	46.6	3.8	470	4	US-09-020-956-102	Sequence 102, App
c	22	46.6	3.8	470	4	US-09-030-607-102	Sequence 102, App
c	23	46.6	3.8	470	4	US-09-439-313-102	Sequence 102, App
	24	46.6	3.8	2277	1	US-08-676-967-2	Sequence 2, Appli
	25	46.6	3.8	2277	1	US-08-676-974-2	Sequence 2, Appli
	26	46.6	3.8	2277	2	US-09-098-487-2	Sequence 2, Appli
	27	46.2	3.8	2295	1	US-08-375-300-3	Sequence 3, Appli
	28	46.2	3.8	2295	3	US-09-177-431-3	Sequence 3, Appli
	29	46.2	3.8	2295	5	PCT-US95-16930-3	Sequence 3, Appli
	30	46.2	3.8	4080	1	US-08-375-300-1	Sequence 1, Appli
	31	46.2	3.8	4080	3	US-09-177-431-1	Sequence 1, Appli
	32	46.2	3.8	4080	5	PCT-US95-16930-1	Sequence 1, Appli
	33	44.2	3.6	696	4	US-09-461-697-193	Sequence 193, App
	34	44.2	3.6	699	4	US-09-461-697-191	Sequence 191, App
	35	44.2	3.6	717	4	US-09-461-697-189	Sequence 189, App
	36	44.2	3.6	774	4	US-09-461-697-187	Sequence 187, App
	37	44.2	3.6	819	4	US-09-461-697-185	Sequence 185, App
	38	44.2	3.6	1669	4	US-09-461-697-184	Sequence 184, App
	39	44	3.6	3207	1	US-07-946-497-1	Sequence 1, Appli
	40	44	3.6	3207	1	US-08-483-322-1	Sequence 1, Appli
	41	44	3.6	3207	2	US-08-478-882-1	Sequence 1, Appli
	42	44	3.6	6638	2	US-08-070-301-2	Sequence 2, Appli
c	43	43.6	3.6	234	1	US-08-469-802B-3	Sequence 3, Appli
c	44	43.6	3.6	234	2	US-08-267-803B-3	Sequence 3, Appli
	45	43.6	3.6	2550	6	5258287-23	Patent No. 5258287

ALIGNMENTS

RESULT 1
US-09-282-305-17
; Sequence 17, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
; APPLICANT: Baldwin, Donald A.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Crane, Virginia C.
; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
; FILE REFERENCE: 5718-44,
; CURRENT APPLICATION NUMBER: US/09/282,305
; CURRENT FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 60/080,563
; PRIOR FILING DATE: 1998-04-03
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 1307
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (87)..(944)
US-09-282-305-17

Query Match 9.2%; Score 112; DB 4; Length 1307;
Best Local Similarity 54.0%; Pred. No. 8.7e-19;
Matches 282; Conservative 0; Mismatches 225; Indels 15; Gaps 2;

```
Qy      59 caatggagttctggggagttgcggtgacaccaaaaaacgctactaaggtgactcctgaag 118
           ||||||||||||||||| | | |||| ||| | | | |     |    || |
Db      85 caatggagttctggggtgaagaagtgaagccaggagccacggtttcttgcaaagttggtg 144

Qy     119 aagacagccttgtccacatttctcaggcttcacttgactgcacagtgaaatctggagaaat 178
           | |      || ||||| |||| |||||| | || |      ||  |||| |  | ||
Db     145 atggtttggttatccacctttcacaggctgccctaggggaaccaaagaaagcgagtgaga 204

Qy     179 ctgtggtttttgagtgtgactgttggtggggctaaacttgttattggaacactttcacaag 238
             ||  || ||  ||| |   ||| ||     ||||| ||  |||||||  | |||  |
Db     205 atgccattgtgtctgtcaaaattgatgataagaaactagtgcttggaaccttatcagttg 264

Qy     239 acaagttccctcagattagctttgatttggtttttgataaagagtttgagctttcacaca 298
           | |||   ||||| ||    | |||| |||| ||||||||||| |||||| | |||||||
Db     265 agaagcatcctcaaatctcttgtgatctggtatttgataaagattttgagttatcacaca 324

Qy     299 gcggtaccaaaagcaaatgttcatttcattggctacaaatcccccaacatcgagcagga-- 356
             | | ||| | ||||   |||  |||||||||| || ||         |   ||
Db     325 attcaaagacagctagtgtttttcttctgtggctacaagtcacctgttcctctgtttgagt 384
```

PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 7.4%; Score 90; DB 21; Length 741;
Best Local Similarity 54.6%; Pred. No. 7.7e-11;
Matches 202; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

```
Qy      55 acaacaatggagttctggggagttgcggtgacaccaaaaaacgctactaaggtgactcct 114
           ||||||||||||||||||||| ||||  || |        ||  |   | |  |   ||||
Db     112 acaacaatggagttctggggtgttgaagttaagaatggcaaaccacttcactagatcct 171

Qy     115 gaagaagacagccttgtccacatttctcaggcttcacttgactgcacagtgaaatctgga 174
           |    |||||| || || ||||| ||||||| | |  | |     |    |||   || |
Db     172 gggctagacaggctcgtgcacatctctcaggttgctttgggagagagtaagaacaatgta 231

Qy     175 ---gaatctgtggttttgagtgtgactgttggtggggctaaacttgttattggaacactt 231
              ||| |  |     |    ||||||||||||   || ||| ||| | ||||||||| ||
Db     232 accgaaccaatccagctttacgtgactgttggatcggataagcttctcattggaacgcta 291

Qy     232 tcacaagacaagttccctcagattagctttgatttggtttttgataaagagtttgagctt 291
           || || || ||||| |||||| |       ||  | || || || |   | |||| |||
Db     292 tctcatgagaagtttcctcagctgtctacggagattgtcttggaaaggaactttgcgctg 351

Qy     292 tcacacagcggtaccaaagcaaatgttcatttcattggctacaaatcccccaacatcgag 351
           || || |   | |  || |  |  |||   |||  ||| ||||||            ||
Db     352 tctcatacttggaagaatgggagcgttttcttctctggttacaaagttgatgcatctgat 411

Qy     352 caggatgacttcactagttcggatgatgaggatgttcctgaagctgttcctgctcctgcc 411
           |  ||         |     |||||| |   ||   |||  | | |    ||| || |
Db     412 cccgagcctgaggatttgattgatgatcaacttgaggctgctggtttcaaagctgctccg 471

Qy     412 cctactgctg 421
           ||||||
Db     472 aaatctgctg 481
```

Search completed: May 2, 2002, 20:26:33
Job time: 8993 sec

PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.

PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 9.1%; Score 110.4; DB 21; Length 331;
Best Local Similarity 95.0%; Pred. No. 1.5e-15;
Matches 114; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy    368 gttcggatgatgaggatgttcctgaagctgttcctgctcctgcccctactgctgttactg 427
          | |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    212 gatcggattatgaggatgttcctgaagctgttcctgctcctgcccctactgctgttactg 271

Qy    428 ccaacggaaatgctggagcagctgttgtcaaggctgacacaaagccaaaggccaaacctg 487
          |||||||||||||||||||||||||||||||||||| ||||||| ||||||| ||||
Db    272 ccaacggaaatgctggagcagctgttgtcaaggctgactcaaagcccaaggccaggcctg 331
```

RESULT 15
AAC40930
ID AAC40930 standard; DNA; 741 BP.
XX
AC AAC40930;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 30042.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.

```
Qy     181 gtggttttgagtgtgactgttggtggggctaaacttgttattggaacactttcacaagac 240
           |   |  ||  ||| |   ||| ||     |||||||  |||||||| || ||    |||
Db     196 gccttgatgtatgtcaaaattgatgatcagaaacttgccattggaaccctctctgttgac 255

Qy     241 aagttccctcagattagctttgatttggtttttgataaagagtttgagctttcacacagc 300
           |||  ||| || ||    |||||| || |||| ||||||||||||||||| || ||||
Db     256 aagaacccacacatacaatttgatctgattttcgataaagagtttgagctgtcgcacaca 315

Qy     301 ggtaccaaagcaaatgttcatttcattggctacaaatccccccaacatcgagcaggatgac 360
             |   |  | |  |||   |||| |||||||||        |        |||| ||
Db     316 tcaaaaactaccagcgttttcttcactggctacaaggttgagcagccattcgaggaagat 375

Qy     361 ttcactagttcggatgatgaggatgttcctgaagctgttcctgctcctgcccctactgct 420
             |    |  |||  ||| ||||     || |   |   || ||         | |
Db     376 gaaatggatcttgattctgaagatgaagacgaggagctgaatgttc---------cagta 426

Qy     421 gttactgccaacggaaatgctggagcagctgttgtcaaggctgacacaaagccaaaggcc 480
           || |  |  || || || ||||  |           ||   | |   |||| ||   ||
Db     427 gtcaaggaaaatggcaaagctgatgagaagaaacagaaaagtcaagaaaaggcagttgct 486

Qy     481 aaacctgccgaagtgaagcctgcagaagagaagcctgaatcagac---------gaggaa 531
             |||| |  ||   |  || |      |||||    |    |||         || ||
Db     487 gcaccttcaaaatcaagtccggattccaagaagagcaaggatgacgacgattctgatgag 546

Qy     532 gatgagtctgatgatgaagatgaagtctgaagaggatgatgactcttgagaaaggaatgg 591
           || ||| ||||||||   ||||| |     |  |      |     | |   |
Db     547 gacgagactgatgattctgatgaggatgagacggacgattctgatgagggtttgtctcct 606

Qy     592 atggttgatgaagatgactcagatgatgacgaggaggaaggattctgaggatgaagaaga 651
           |  |  || ||||| |||  |||||| || ||     |   |   |       |
Db     607 gaagaaggcgatgatgattcaagtgatgaagatgataccagtgacgatgaggaggaagac 666

Qy     652 ggaggagctcttaagaagcctgagccaatcaacaagaagaggccaaatgaatctgtatcc 711
             | ||| ||||||||||||||  |  ||| ||||     |  |    || ||
Db     667 actccaactcctaagaagcctgaggtaggcaagaagagagctgctgaaagttccgtgctg 726

Qy     712 aaaacacccgtctctggaaagaaggcaaaaccagcagcagcaccagct 759
           ||||| ||  | ||||  ||||| |||||    ||  || |  || ||
Db     727 aaaactcctctatctgataagaaagcaaaggttgccacaccgtcatct 774


RESULT 14
AAC52921
ID   AAC52921 standard; DNA; 331 BP.
XX
AC   AAC52921;
XX
DT   18-OCT-2000 (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 72846.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
```

```
PT   New deacetylase genes, used for producing transgenic plants which have
PT   increased disease resistance  -
XX
PS   Claim 1; Page 77-79; 87pp; English.
XX
CC   The present sequence encodes a maize histone deacetylase . This DNA
CC   belongs to family 2, ZmHD2 and appears to affect chromatin structure at
CC   promoters of RNA polymerase I and thus regulate ribosomal RNA production.
CC   The nucleotide sequence can be used to transform plants and increase
CC   disease resistance by altering the levels of heterochromatin, altering
CC   chromatin assembly, and gene activity of the transformed plants.
CC   Additionally, compositions find use in screening for toxins that affect
CC   pathogenicity and in determining which disease response promoters are
CC   regulated by histone deacetylase.
XX
SQ   Sequence 1245 BP; 385 A; 250 C; 324 G; 286 T; 0 other;


  Query Match             9.1%;  Score 111.2;  DB 20;  Length 1245;
  Best Local Similarity   51.6%;  Pred. No. 1.5e-15;
  Matches  430;  Conservative    0;  Mismatches  363;  Indels   41;  Gaps    6;

Qy      13 taaaaaaaaacctaacaacctctcttctctcttcctcgttcaacaacaatggagttctgg 72
           | |||     |  |  ||  |  |||||  |     |||    |  |  | ||||||||||||
Db      36 ttaaaccctagttcccattccatcttcgtttccgctccgccgccgtcgatggagttctgg 95

Qy      73 ggagttgcggtgacaccaaaaaacgctactaaggtgactcctgaagaagacagccttgtc 132
           ||  | | ||| | |||   |  | ||   |||      |||| | | | |    |  |
Db      96 ggtctcgaggtcaaacctggatccactgtcaagtgtgagcctggacatggctttatcctg 155

Qy     133 cacatttctcaggcttcacttgactgcacagtgaaatctggagaatctgtggttttgagt 192
           ||| |||| |||||| | ||||      ||  ||||  |   ||   ||   |  ||  |
Db     156 cacgtttcccaggctgcccttggggaatcaaagaaaagt---gacagtgccttaatgtat 212

Qy     193 gtgactgttggtggggctaaacttgttattggaacactttcacaagacaagttccctcag 252
           || |  |||| ||     || ||||  |||||||| || ||    ||||| | ||| |||
Db     213 gtcaaagttgatgacaagaagcttgccattggaacgctctctatcgacaaatacccacag 272

Qy     253 attagctttgatttggtttttgataaagagtttgagctttcacacagcggtaccaaagca 312
           ||    || ||||||||||| ||||||||||||||||| |||||||     |  |   |
Db     273 atacaattcgatttggttttcaataaagagtttgagctgtcacacacatcgaaaactacc 332

Qy     313 aatgttcatttcattggctacaaatcccccaacatcgagcaggatgacttcactagttcg 372
           | |||   ||||  ||| |||||         |       ||| ||    |    |
Db     333 agtgtatttttctctggttacaaggttgagcagccaattgagggagatgaaatggatctt 392

Qy     373 gatgatgaggatgttcctgaagctgttcctgctcctg-----------ccctactgct 420
           |||  ||||||||     || |   |      ||| |                |  |||
Db     393 gattctgaggatgaagaggaggagctaaacattccagtaatcaaggaaaatggcaaagct 452

Qy     421 gttactgccaacggaaatgctggagcagctgttgtcaaggctgacacaaagccaaaggcc 480
           | |       | |   |          ||     |    ||||  |||    ||||  |
Db     453 gatgggaaggaggagcagaaaaatcaagagaaggcagtagctgctacagcttcaaaatca 512

Qy     481 aaacctgccgaagtgaagcctgcagaagagaagcctgaatcagacgaggaagatgagtct 540
           |  | || |   |  |||       | ||  |  ||||       || || |||||||||
Db     513 agtcttggccttgaaaagaaaagcaaggatgactctgatgattctgatgaggatgagtct 572

Qy     541 gatgatgaagatgaagtctgaagaggatgatgactcttgagaaaggaatggat------- 593
           ||||||   |||||     || ||   |||||| |||   ||| |  | ||||
Db     573 gatgattctgatga----ggatgattctgatgattctgatgaaggcgagggattatctcc 628

Qy     594 ---ggttgatgaagatgactcagatgatgacgaggaggaaggattctgaggatgaagaag 650
               |  | || ||||| |||  |||||| || ||     |       |       |
Db     629 tgacgaaggcgatgatgattcaagtgatgaggatgataccagtgatgatgacgaggaaga 688

Qy     651 aggaggagctcttaagaagcctgagccaatcaacaagaagaggccaaatgaatctgtatc 710
           |     | ||| ||| ||||| ||| ||     |||||||||      |||| ||
Db     689 aaccccaactcctaaaaagccagaggca---ggcaagaagagaggtgctgaaaatgctct 745

Qy     711 caaaacacccgtctctggaaagaaggcaaaaccagcagcagcaccagcttctactcctca 770
           |||||  ||  | ||||  |||||  |||||     ||  || | ||||          | ||
Db     746 gaaaacgcctctttctgataagaaagcaaaggttgccacaccgccag---------ccca 796
```

```
Qy     771 gaagacagaagagaagaagaaaggaggacacaccgccacaccacacccagctaa 824
           ||| ||||  |  ||||||   |     ||    || || ||||||||||| ||
Db     797 gaaaacaggtggcaagaagggtgccacccatgtggcaactccacacccagcaaa 850


RESULT  13
AAX90842
ID   AAX90842 standard; DNA; 1283 BP.
XX
AC   AAX90842;
XX
DT   13-JAN-2000  (first entry)
XX
DE   Maize histone deacetylase-6 DNA.
XX
KW   Maize histone deacetylase; family 2, ZmHD2; chromatin structure;
KW   RNA polymerase I; ribosomal RNA production; promoter regulator; promoter;
KW   transcription; plant transformation; heterochromatin; disease resistance;
KW   chromatin assembly; gene activity; toxin screening; pathogenicity;
KW   disease response promoter; ds.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   CDS             38..1336
FT                   /*tag= a
FT                   /product= "Maize histone deacetylase"
XX
PN   WO9951731-A2.
XX
PD   14-OCT-1999.
XX
PF   02-APR-1999;   99WO-US07370.
XX
PR   03-APR-1998;   98US-0080563.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Baldwin DA,  Briggs SP,  Crane VC;
XX
DR   WPI; 1999-611038/52.
DR   P-PSDB; AAY28802.
XX
PT   New deacetylase genes, used for producing transgenic plants which have
PT   increased disease resistance  -
XX
PS   Claim 1; Page 70-72; 87pp; English.
XX
CC   The present sequence encodes a maize histone deacetylase . This DNA
CC   belongs to family 2, ZmHD2 and appears to affect chromatin structure at
CC   promoters of RNA polymerase I and thus regulate ribosomal RNA production.
CC   The nucleotide sequence can be used to transform plants and increase
CC   disease resistance by altering the levels of heterochromatin, altering
CC   chromatin assembly, and gene activity of the transformed plants.
CC   Additionally, compositions find use in screening for toxins that affect
CC   pathogenicity and in determining which disease response promoters are
CC   regulated by histone deacetylase.
XX
SQ   Sequence 1283 BP; 383 A; 259 C; 333 G; 308 T; 0 other;


  Query Match             9.1%;  Score 111.2;  DB 20;  Length 1283;
  Best Local Similarity   50.7%;  Pred. No. 1.5e-15;
  Matches  359;  Conservative    0;  Mismatches  328;  Indels   21;  Gaps    3;

Qy      61 atggagttctggggagttgcggtgacaccaaaaaacgctactaaggtgactcctgaagaa 120
           ||||||||||||||  | | ||| |  ||       | ||  ||||      |||| | |
Db      79 atggagttctggggtctcgaggtcaagcctggttccactgttaagtgtgaggcctggatat 138

Qy     121 gacagccttgtccacatttctcaggcttcacttgactgcacagtgaaatctggagaatct 180
           | |    |  | ||| |||| ||||||| | ||||   |   |  |||     | ||   |
Db     139 ggctttgtgctgcacctttcccaggctgctcttg---gggaatcgaagaagagtgataat 195
```

```
AAX90843
ID   AAX90843 standard; DNA; 1191 BP.
XX
AC   AAX90843;
XX
DT   13-JAN-2000  (first entry)
XX
DE   Maize histone deacetylase-7 DNA.
XX
KW   Maize histone deacetylase; family 2, ZmHD2; chromatin structure;
KW   RNA polymerase I; ribosomal RNA production; promoter regulator; promoter;
KW   transcription; plant transformation; heterochromatin; disease resistance;
KW   chromatin assembly; gene activity; toxin screening; pathogenicity;
KW   disease response promoter; ds.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   CDS             63..971
FT                   /*tag=  a
FT                   /product= "Maize histone deacetylase"
XX
PN   WO9951731-A2.
XX
PD   14-OCT-1999.
XX
PF   02-APR-1999;   99WO-US07370.
XX
PR   03-APR-1998;   98US-0080563.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Baldwin DA,  Briggs SP,  Crane VC;
XX
DR   WPI; 1999-611038/52.
DR   P-PSDB; AAY28803.
XX
PT   New deacetylase genes, used for producing transgenic plants which have
PT   increased disease resistance  -
XX
PS   Claim 1; Page 73-75; 87pp; English.
XX
CC   The present sequence encodes a maize histone deacetylase . This DNA
CC   belongs to family 2, ZmHD2 and appears to affect chromatin structure at
CC   promoters of RNA polymerase I and thus regulate ribosomal RNA production.
CC   The nucleotide sequence can be used to transform plants and increase
CC   disease resistance by altering the levels of heterochromatin, altering
CC   chromatin assembly, and gene activity of the transformed plants.
CC   Additionally, compositions find use in screening for toxins that affect
CC   pathogenicity and in determining which disease response promoters are
CC   regulated by histone deacetylase.
XX
SQ   Sequence 1191 BP; 365 A; 242 C; 311 G; 273 T; 0 other;


  Query Match             9.2%;  Score 111.6;  DB 20;  Length 1191;
  Best Local Similarity   51.4%;  Pred. No. 1.2e-15;
  Matches  362;  Conservative    0;  Mismatches  329;  Indels   13;  Gaps    4;

Qy      61 atggagttctggggagttgcggtgacaccaaaaaacgctactaaggtgactcctgaagaa 120
           ||||||||||||||  | | ||| |  ||      | ||  ||||      |||| | |
Db      63 atggagttctggggtctcgaggtcaagcctggttccactgttaagtgtgagcctggatat 122

Qy     121 gacagccttgtccacatttctcaggcttcacttgactgcacagtgaaatctggagaatct 180
           | |    |  | ||| |||| |||||| | ||||   |   |  |||     | ||   |
Db     123 ggctttgtgctgcacctttcccaggctgctcttg---gggaatcgaagaagagtgataat 179

Qy     181 gtggttttgagtgtgactgttggtggggctaaacttgttattggaacactttcacaagac 240
           |   |  ||  ||| |   ||| ||     |||||||  |||||||| || ||    |||
Db     180 gccttgatgtatgtcaaaattgatgatcagaaacttgccattggaaccctctctgttgac 239

Qy     241 aagttccctcagattagctttgatttggtttttgataaagagtttgagctttcacacagc 300
           |||  ||| || |||   |||||| || |||| ||||||||||||||||| || ||||

Db     240 aagaacccacacattcaatttgatctgattttcgataaagagtttgagctgtcgcacaca 299

Qy     301 ggtaccaaagcaaatgttcatttcattggctacaaatcccccaacatcgagcaggatgac 360
           |  |   | |  ||    |||| |||||||||        |       |||| ||
Db     300 tcaaaaactaccagcgtcttcttcactggctacaaggttgaacagccattcgaggaagat 359

Qy     361 ttcactagttcggatgatgaggatgttcctgaagctgttcctgctcctgcccctactg-- 418
           |    |  |||  ||| ||||     || |   |   || ||| |     |  |
Db     360 gaaatggatcttgattctgaagatgaagacgaggagctgaatgttccagtagtcaaggaa 419

Qy     419 -ctgttactgccaacggaaatgctggagcagctgttgtcaaggctgacacaaagccaaag 477
            ||  |  ||  | || ||         |  |   |  ||||| |   |    ||
Db     420 aatggcaaagctgatgggaagaaacagaaaagtcaagaaaaggcagttgctgcaccttca 479

Qy     478 gccaaacctgccgaagtgaagcctgcagaagagaagcctgaatcagacgaggaagatgag 537
           |  | | ||    |||       | ||  |   ||| || || ||||| ||
Db     480 aaatcaagtccggattccaagaagagcaaggatgacgatgattctgatgaggacgagact 539

Qy     538 tctgatgatgaagatga--agtctgaagaggatgatgactcttgagaaaggaatggatgg 595
           ||||  ||| || ||  || | || ||   ||||||  || |       | | |
Db     540 gatgattctgatgaggatgagacggacgattctgatgagggtttatcttctgaagaa--- 596

Qy     596 ttgatgaagatgactcagatgatgacgaggaggaaggattctgaggatgaagaagaggag 655
           |  || ||||| ||| |||||| || ||      |   |   |       |
Db     597 --ggcgatgatgattcaagtgatgaagatgataccagtgacgatgaggaggaagacactc 654

Qy     656 gagctcttaagaagcctgagccaatcaacaagaagaggccaaatgaatctgtatccaaaa 715
           | ||| ||||||||||||  |  ||| ||||      |  |    || ||    ||||
Db     655 caactcctaagaagcctgaggtaggcaagaagagagctgctgaaagttccgtgctgaaaa 714

Qy     716 cacccgtctctggaaagaaggcaaaaccagcagcagcaccagct 759
           | ||  | ||||  ||||| |||||    ||  || |  || ||
Db     715 ctcctctatctgataagaaagcaaaggttgccacaccgtcatct 758


RESULT  12
AAX90844
ID   AAX90844 standard; DNA; 1245 BP.
XX
AC   AAX90844;
XX
DT   13-JAN-2000  (first entry)
XX
DE   Maize histone deacetylase-8 DNA.
XX
KW   Maize histone deacetylase; family 2, ZmHD2; chromatin structure;
KW   RNA polymerase I; ribosomal RNA production; promoter regulator; promoter;
KW   transcription; plant transformation; heterochromatin; disease resistance;
KW   chromatin assembly; gene activity; toxin screening; pathogenicity;
KW   disease response promoter; ds.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   CDS             84..1019
FT                   /*tag=  a
FT                   /product= "Maize histone deacetylase"
XX
PN   WO9951731-A2.
XX
PD   14-OCT-1999.
XX
PF   02-APR-1999;   99WO-US07370.
XX
PR   03-APR-1998;   98US-0080563.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Baldwin DA,  Briggs SP,  Crane VC;
XX
DR   WPI; 1999-611038/52.
DR   P-PSDB; AAY28804.
XX
```

```
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match           12.3%;  Score 149.8;  DB 21;  Length 774;
  Best Local Similarity   64.8%;  Pred. No. 2.5e-24;
  Matches  239;  Conservative    0;  Mismatches  127;  Indels    3;  Gaps    1;

Qy      61 atggagttctggggagttgcggtgacaccaaaaaacgctactaaggtgactcctgaagaa 120
           ||||||||||||||| |||  || | | ||  ||| |   ||  |||||||||||||||
Db       1 atggagttctggggaattgaagttaaatcaggaaagccagttacagtgactcctgaagaa 60

Qy     121 gacagccttgtccacatttctcaggcttcactt---gactgcacagtgaaatctggagaa 177
           | ||  ||| ||||| |||||||||| || |||   || || | |   ||    |||||
Db      61 ggcattcttatccacgtttctcaggcatcgcttggagaatgtaaaaacaagaagggagag 120

Qy     178 tctgtggttttgagtgtgactgttggtggggctaaacttgttattggaacactttcacaa 237
           | ||||  |||   ||| |  |||||       ||  | ||| | ||||| || ||
Db     121 tttgtgcctttacatgtaaaggttgggaaccagaacttggttctgggaactctatcgact 180

Qy     238 gacaagttccctcagattagctttgatttggtttttgataaagagtttgagctttcacac 297
           || ||  |||||||| ||  || |||||||| || || || ||||||||||||| |||
Db     181 gagaacatccctcagctttctgtgatttggtattcgacaaggagtttgagctttctcac 240

Qy     298 agcggtaccaaagcaaatgttcatttcattggctacaaatccccccaacatcgagcaggat 357
           |   |    |||| || |||| | ||  |||| |||||| | |||||||| ||||   |
Db     241 acttggggaaaaggaagtgtttactttgttggatacaaaactcccaacattgagccacaa 300

Qy     358 gacttcactagttcggatgatgaggatgttcctgaagctgttcctgctcctgcccctact 417
           | ||   ||      || || || || |     |||| | | |||    |||  | |
Db     301 ggctattctgaggaagaagaggaagaagaggaagaagttcctgctgggaatgctgccaag 360

Qy     418 gctgttact 426
           |||||  ||
Db     361 gctgtagct 369


RESULT  10
AAX90845
ID   AAX90845 standard; DNA; 1307 BP.
XX
AC   AAX90845;
XX
DT   13-JAN-2000  (first entry)
XX
DE   Maize histone deacetylase-9 DNA.
XX
KW   Maize histone deacetylase; family 2, ZmHD2; chromatin structure;
KW   RNA polymerase I; ribosomal RNA production; promoter regulator; promoter;
KW   transcription; plant transformation; heterochromatin; disease resistance;
KW   chromatin assembly; gene activity; toxin screening; pathogenicity;
KW   disease response promoter; ds.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   CDS             87..944
FT                   /*tag= a
FT                   /product= "Maize histone deacetylase"
XX
PN   WO9951731-A2.
XX
PD   14-OCT-1999.
XX
PF   02-APR-1999;   99WO-US07370.
XX
PR   03-APR-1998;   98US-0080563.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Baldwin DA,  Briggs SP,  Crane VC;
XX
DR   WPI; 1999-611038/52.
DR   P-PSDB; AAY28805.
XX
PT   New deacetylase genes, used for producing transgenic plants which have
PT   increased disease resistance  -
XX
PS   Claim 1; Page 80-82; 87pp; English.
XX
CC   The present sequence encodes a maize histone deacetylase . This DNA
CC   belongs to family 2, ZmHD2 and appears to affect chromatin structure at
CC   promoters of RNA polymerase I and thus regulate ribosomal RNA production.
CC   The nucleotide sequence can be used to transform plants and increase
CC   disease resistance by altering the levels of heterochromatin, altering
CC   chromatin assembly, and gene activity of the transformed plants.
CC   Additionally, compositions find use in screening for toxins that affect
CC   pathogenicity and in determining which disease response promoters are
CC   regulated by histone deacetylase.
XX
SQ   Sequence 1307 BP; 386 A; 260 C; 308 G; 353 T; 0 other;


  Query Match           9.2%;  Score 112;  DB 20;  Length 1307;
  Best Local Similarity   54.0%;  Pred. No. 9.9e-16;
  Matches  282;  Conservative    0;  Mismatches  225;  Indels   15;  Gaps    2;

Qy      59 caatggagttctggggagttgcggtgacaccaaaaaacgctactaaggtgactcctgaag 118
           |||||||||||||||| |  |  |||| ||| |  | |   |      |    ||  |
Db      85 caatggagttctggggtgaagaagtgaagccaggagccacggtttcttgcaaagttggtg 144

Qy     119 aagacagccttgtccacatttctcaggcttcacttgactgcacagtgaaatctggagaat 178
           | |      || ||||| |||| |||||| | || |      ||  |||| |  | ||
Db     145 atggtttggttatccacctttcacaggctgccctaggggaaccaaagaaagcgagtgaga 204

Qy     179 ctgtggttttgagtgtgactgttggtggggctaaacttgttattggaacactttcacaag 238
           ||  || ||  ||| |   ||| ||     ||||| ||  |||||||  | |||   |
Db     205 atgccattgtgtctgtcaaaattgatgataagaaactagtgcttggaaccttatcagttg 264

Qy     239 acaagttccctcagattagctttgatttggtttttgataaagagtttgagctttcacaca 298
           | |||   ||||| ||    | |||| |||| |||||||||| |||||| | |||||||
Db     265 agaagcatcctcaaatctcttgtgatctggtatttgataaagattttgagttatcacaca 324

Qy     299 gcggtaccaaagcaaatgttcatttcattggctacaaatcccccaacatcgagcagga-- 356
               |  | ||| | ||||   |||  |||||||||| || ||          |   ||
Db     325 attcaaagacagctagtgttttcttctgtggctacaagtcacctgttcctctgtttgagt 384

Qy     357 -------tgacttcactagttcggatgatgaggatgttcctgaagctgttcctgctcc-- 407
                  ||       |||||| ||||| |||| ||  || ||     ||||    |
Db     385 ctgattctggtgaagatagttcagatgaagaggttgaacccgatctaattccaatgcaga 444

Qy     408 ----tgcccctactgctgttactgccaacggaaatgctggagcagctgttgtcaaggctg 463
               ||    ||    |  |||||| || |     |     |  | | |    || ||||
Db     445 ataatgaaattaaaatttctactgcaaaggttcccgtgaaggttggtatacaaaatgctg 504

Qy     464 acacaaagccaaaggccaaacctgccgaagtgaagcctgcagaagagaagcctgaatcag 523
           |   | |   ||   | |    ||  || | |     |    | ||     |   ||||
Db     505 atgaagatgaaacttctagtggtgatgatgatgatttcactgatagtgatagtgaaatgt 564

Qy     524 acgaggaagatgagtctgatgatgaagatgaagtctgaagag 565
           |||||||||||||||   |||||||||||||| | ||| |
Db     565 ctgaggaagatgagtccagtgatgaagatgaagtgtcaagcg 606


RESULT  11
```

PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.

```
Db    302 cttggggaaaaggaagtgtttactttgttggatacaaaactcccaacattgagccacaag 361

Qy    359 acttcactagttcggatgatgaggatgttcctgaagctgttcctgctcctgcccctactg 418
Db    362 gctattctgaggaagaagaggaagaagaggaagaagttcctgctgggaatgctgccaagg 421

Qy    419 ctgttact 426
Db    422 ctgtagct 429


RESULT   8
AAF80352
ID   AAF80352 standard; DNA; 939 BP.
XX
AC   AAF80352;
XX
DT   29-JUN-2001  (first entry)
XX
DE   Nucleotide sequence of a histone deacetylase designated AtHDA2A.
XX
KW   Histone deacetylase; AtHDA2A; RPD3; gene expression; transgenic plant;
KW   HDA1; ethylene-responsive phenotype; hypocotyl elongation; ds.
XX
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   CDS             49..786
FT                   /*tag= a
FT                   /product= "histone deacetylase AtHDA2A"
XX
PN   CA2316036-A1.
XX
PD   27-FEB-2001.
XX
PF   24-AUG-2000; 2000CA-2316036.
XX
PR   27-AUG-1999;   99US-0383971.
XX
PA   (MIAC ) CANADA MIN AGRIC & AGRI-FOOD CANADA.
XX
PI   Miki B,  Brown D,  Tian L,  Wu K;
XX
DR   WPI; 2001-258457/27.
DR   P-PSDB; AAB67813.
XX
PT   Methods for regulating gene expression in transgenic plants, e.g.
PT   repressing ethylene-responsive phenotypes (e.g. inhibition of hypocotyl
PT   elongation), comprises introducing genes encoding histone deacetylase
PT   -
XX
PS   Claim 9; Fig 2A; 91pp; English.
XX
CC   The present sequence encodes an Arabidopsis thaliana histone deacetylase
CC   designated AtHDA2A. The protein is homologous to yeast RPD3 and HDA1.
CC   The polynucleotide sequence is used in the method of the invention.
CC   The specification describes a method for regulating gene expression in
CC   transgenic plants. The method comprises modifying histones by introducing
CC   chimeric nucleotide sequences which have regulatory elements in operative
CC   association with a gene of interest or with a nucleotide sequence
CC   encoding histone deacetylase. The method is useful for regulating the
CC   developmental, physiological or biochemical pathway within a plant,
CC   particularly for repressing ethylene-responsive phenotypes
CC   (e.g. inhibition of hypocotyl elongation). The method is also useful
CC   as a functional test for identifying a phenotype associated with
CC   perturbing a gene. The histone deacetylase genes are useful for
CC   altering the development of an organism.
XX
SQ   Sequence 939 BP; 296 A; 175 C; 236 G; 232 T; 0 other;

  Query Match             12.8%;  Score 156.2;  DB 22;  Length 939;
  Best Local Similarity   64.1%;  Pred. No. 9.3e-26;
  Matches  252;  Conservative    0;  Mismatches  138;  Indels    3;  Gaps    1;

Qy     37 ttctctcttcctcgttcaacaacaatggagttctggggagttgcggtgacaccaaaaaac 96
Db     25 ttttctcaaccttgattcttagccatggagttctggggaattgaagttaaatcaggaaag 84

Qy     97 gctactaaggtgactcctgaagaagacagccttgtccacatttctcaggcttcactt--- 153
Db     85 ccagttacagtgactcctgaagaaggcattcttatccacgtttctcaggcatcgcttgga 144

Qy    154 gactgcacagtgaaatctggagaatctgtggttttgagtgtgactgttggtggggctaaa 213
Db    145 gaatgtaaaaacaagaagggagagtttgtgccttacatgtaaaggttgggaaccagaac 204

Qy    214 cttgttattggaacactttcacaagacaagttccctcagattagctttgatttggttttt 273
Db    205 ttggttctgggaactctatcgactgagaacatccctcagcttttctgtgatttggtattc 264

Qy    274 gataaagagtttgagctttcacacagcggtaccaaagcaaatgttcatttcattggctac 333
Db    265 gacaaggagtttgagctttctcacacttggggaaaaggaagtgtttactttgttggatac 324

Qy    334 aaatcccccaacatcgagcaggatgacttcactagttcggatgatgaggatgttcctgaa 393
Db    325 aaaactcccaacattgagccacaaggctattctgaggaagaagaggaagaagaggaagaa 384

Qy    394 gctgttcctgctcctgcccctactgctgttact 426
Db    385 gttcctgctgggaatgctgccaaggctgtagct 417


RESULT   9
AAC43043
ID   AAC43043 standard; DNA; 774 BP.
XX
AC   AAC43043;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 37816.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
```

PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.

PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 13.1%; Score 159.2; DB 21; Length 991;
Best Local Similarity 62.4%; Pred. No. 2e-26;
Matches 267; Conservative 0; Mismatches 158; Indels 3; Gaps 1;

Qy 2 tctttcgcttctaaaaaaaaacctaacaacctctcttctctcttcctcgttcaacaacaa 61
Db 2 ttttcagctgctcataaaccctaaaaatcctctctttttctcaaccttgattcttagcca 61

Qy 62 tggagttctggggagttgcggtgacaccaaaaaacgctactaaggtgactcctgaagaag 121
Db 62 tggagttctggggaattgaagttaaatcaggaaagccagttacagtgactcctgaagaag 121

Qy 122 acagccttgtccacatttctcaggcttcactt---gactgcacagtgaaatctggagaat 178
Db 122 gcattcttatccacgtttctcaggcatcgcttggagaatgtaaaaacaagaagggagagt 181

Qy 179 ctgtggttttgagtgtgactgttggtggggctaaacttgttattggaacactttcacaag 238
Db 182 ttgtgcctttacatgtaaaggttgggaaccagaacttggttctgggaactctatcgactg 241

Qy 239 acaagttccctcagattagctttgatttggtttttgataaagagtttgagctttcacaca 298
Db 242 agaacatccctcagcttttctgtgatttggtattcgacaaggagtttgagctttctcaca 301

Qy 299 gcggtaccaaagcaaatgttcatttcattggctacaaatcccccaacatcgagcaggatg 358

```
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match             13.3%;  Score 162.6;  DB 21;  Length 1002;
  Best Local Similarity   57.1%;  Pred. No. 3.4e-27;
  Matches  316;  Conservative    0;  Mismatches  234;  Indels    3;  Gaps    1;

Qy       2 tctttcgcttctaaaaaaaaacctaacaacctctcttctctcttcctcgttcaacaacaa 61
Db       4 ttttcagctgctcataaaccctaaaaatcctctctttttctcaaccttgattcttagcca 63

Qy      62 tggagttctggggagttgcggtgacaccaaaaaacgctactaaggtgactcctgaagaag 121
Db      64 tggagttctggggaattgaagttaaatcaggaaagccagttacagtgactcctgaagaag 123

Qy     122 acagccttgtccacatttctcaggcttcactt---gactgcacagtgaaatctggagaat 178
Db     124 gcattcttatccacgtttctcaggcatcgcttggagaatgtaaaaacaagaagggagagt 183

Qy     179 ctgtggttttgagtgtgactgttggtggggctaaacttgttattggaacactttcacaag 238
Db     184 ttgtgcctttacatgtaaaggttgggaaccagaacttggttctgggaactctatcgactg 243

Qy     239 acaagttccctcagattagctttgatttggtttttgataaagagtttgagctttcacaca 298
Db     244 agaacatccctcagcttttctgtgatttggtattcgacaaggagtttgagctttctcaca 303

Qy     299 gcggtaccaaagcaaatgttcatttcattggctacaaatcccccaacatcgagcaggatg 358
Db     304 cttggggaaaaggaagtgtttactttgttggatacaaaactcccaacattgagccacaag 363

Qy     359 acttcactagttcggatgatgaggatgttcctgaagctgttcctgctcctgcccctactg 418
Db     364 gctattctgaggaagaagaggaagaagaggaagaagaggaagaagttcctgctgggaatg 423

Qy     419 ctgttactgccaacggaaatgctggagcagctgttgtcaaggctgacacaaagccaaagg 478
Db     424 ctgccaaggctgtagctaaaccaaaggctaagcctgcagaagtgaagccagctgttgatg 483

Qy     479 ccaaacctgccgaagtgaagcctgcagaagaagaagcctgaatcagacgaggaagatgagt 538
Db     484 atgaagaggatgagtctgattctgacggaatggatgaagatgattctgatggtgaggatt 543

Qy     539 ctgatgatgaaga 551
Db     544 ctgaggaagaaga 556


RESULT   7
AAC45610
ID   AAC45610 standard; DNA; 991 BP.
XX
AC   AAC45610;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 47120.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
```

PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.

```
PR   16-AUG-1999;   99US-0149368.
PR   17-AUG-1999;   99US-0149175.
PR   18-AUG-1999;   99US-0149426.
PR   20-AUG-1999;   99US-0149722.
PR   20-AUG-1999;   99US-0149723.
PR   20-AUG-1999;   99US-0149929.
PR   23-AUG-1999;   99US-0149902.
PR   23-AUG-1999;   99US-0149930.
PR   25-AUG-1999;   99US-0150566.
PR   26-AUG-1999;   99US-0150884.
PR   27-AUG-1999;   99US-0151065.
PR   27-AUG-1999;   99US-0151066.
PR   27-AUG-1999;   99US-0151080.
PR   30-AUG-1999;   99US-0151303.
PR   31-AUG-1999;   99US-0151438.
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match           13.6%;  Score 165.2;  DB 21;  Length 940;
  Best Local Similarity   58.4%;  Pred. No. 8.7e-28;
  Matches  326;  Conservative    0;  Mismatches  228;  Indels    4;  Gaps    2;

Qy       2 tctttcgcttctaaaaaaaaacctaacaacctctcttctctcttcctcgttcaacaacaa 61
           | ||  ||| || | |||      ||   |     || ||||  ||| | |    | | |
Db       4 ttttcagctgctcataaaccctaaaaatcctctctttttctcaaccttgattcttagcca 63

Qy      62 tggagttctggggaattgcggtgacaccaaaaaacgctactaaggtgactcctgaagaag 121
           ||||||||||||||| |||  || | | ||  |||  |   ||  ||||||||||||||||
Db      64 tggagttctggggaattgaagttaaatcaggaaagccagttacagtgactcctgaagaag 123

Qy     122 acagccttgtccacatttctcaggcttcactt---gactgcacagtgaaatctggagaat 178
           ||  ||| ||||| ||||||||||| || |||   || || | |   ||    ||||| |
Db     124 gcattcttatccacgtttctcaggcatcgcttggagaatgtaaaaacaagaaggggagagt 183

Qy     179 ctgtggttttgagtgtgactgttggtggggctaaacttgttattggaacactttcacaag 238
           ||||  |||   ||| |  |||||        ||  | ||| | ||||| || ||    |
Db     184 ttgtgcctttacatgtaaaggttgggaaccagaacttggttctgggaactctatcgactg 243

Qy     239 acaagttccctcagattagctttgatttggtttttgataaagagtttgagctttcacaca 298
           | ||  |||||||| ||  || ||||||||| || || || |||||| ||||||| ||||
Db     244 agaacatccctcagcttttctgtgatttggtattcgacaaggagttt-agctttctcaca 302

Qy     299 gcggtaccaaagcaaatgttcatttcattggctacaaatcccccaacatcgagcaggatg 358
           |    |||| || |||| | ||  |||| |||||| | |||||||| ||||   | |
Db     303 cttggggaaaaggaagtgtttacttgttggatacaaaactcccaacattgagccacaag 362

Qy     359 acttcactagttcggatgatgaggatgttcctgaagctgttcctgctcctgcccctactg 418
           ||   ||      || || || || |     |||| |  | ||||   | |  ||| ||
Db     363 gctattctgaggaagaagaggaagaagaggaagaagttcctgctgccgttactgctagtg 422

Qy     419 ctgttactgccaacggaaatgctggagcagctgttgtcaaggctgacacaaagccaaagg 478
           | ||||||| |     |||  | ||   | |      || ||    ||||||   |
Db     423 ggaatgctgccaaggctgtagctaaaccaaaggctaagcctgcagaagtgaagccagctg 482

Qy     479 ccaaacctgccgaagtgaagcctgcagaagagaagcctgaatcagacgaggaagatgagt 538
           |   ||  || | || |||     ||       ||   ||| ||    ||||
Db     483 ttgatgatgaagaggatgagtctgattctgacggaatggatgaagatgattctgatggtg 542

Qy     539 ctgatgatgaagatgaag 556
           |||  ||| || ||||
Db     543 aggattctgaggaagaag 560


RESULT   6
AAC35217
ID   AAC35217 standard; DNA; 1002 BP.
XX
AC   AAC35217;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 9400.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
```

```
Qy    740 aaccagcagcagcaccagcttctactcctcagaagac 776
          |||||||||||||||||||||||||||||||||||||
Db    432 aaccagcagcagcaccagcttctactcctcagaagac 468


RESULT   5
AAC36931
ID   AAC36931 standard; DNA; 940 BP.
XX
AC   AAC36931;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 15573.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
PR   02-AUG-1999;   99US-0146386.
PR   02-AUG-1999;   99US-0146388.
PR   02-AUG-1999;   99US-0146389.
PR   03-AUG-1999;   99US-0147038.
PR   04-AUG-1999;   99US-0147204.
PR   04-AUG-1999;   99US-0147302.
PR   05-AUG-1999;   99US-0147192.
PR   05-AUG-1999;   99US-0147260.
PR   06-AUG-1999;   99US-0147303.
PR   06-AUG-1999;   99US-0147416.
PR   09-AUG-1999;   99US-0147493.
PR   09-AUG-1999;   99US-0147935.
PR   10-AUG-1999;   99US-0148171.
PR   11-AUG-1999;   99US-0148319.
PR   12-AUG-1999;   99US-0148341.
PR   13-AUG-1999;   99US-0148565.
PR   13-AUG-1999;   99US-0148684.
```

PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 30.3%; Score 369.4; DB 21; Length 468;
Best Local Similarity 95.0%; Pred. No. 5.4e-74;
Matches 434; Conservative 1; Mismatches 17; Indels 5; Gaps 5;

```
Qy    321 tttcattggctacaaatcccccaacatcgagcaggatgacttcactagttcggatgatga 380
          ||||  |    | |||   || ||||||||||||||||||||||||||||||||||||
Db     16 tttcgcttctaaaaaaaaacctaacatcgagcaggatgacttcactagttcggatgatga 75

Qy    381 ggatgttcctgaagctgttcctgctcctgcccctactgctgttactgccaacggaaatgc 440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     76 ggatgttcctgaagctgttcctgctcctgcccctactgctgttactgccaacggaaatgc 135

Qy    441 tggagcagctgttgtcaaggctgacacaaagccaaaggccaaacctgccgaagtgaagcc 500
          |||||||||||||||||||||||| ||||||| |||||||| ||||||||||||||||
Db    136 tggagcagctgttgtcaaggctgactcaaagcccaaggccaagcctgccgaagtgaagcc 195

Qy    501 tgcagaagagaagcctgaatcagacgaggaagatgagtctgatgatgaagatgaagtctg 560
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||| ||||||
Db    196 tgcagaagagaagccggaatcagacgaggaagatgagtctgatgatgaagatg-agtctg 254

Qy    561 aagaggatgatgactcttgaqaaaqgaatgqatgqttgatgaagatgactcagatgatga 620
          |||||||||||||||| |||||||||||||||| ||||||||||||||||||||||||||
Db    255 aagaggatgatgactc-tgagaaaggaatggat-gttgatgaagatgactcagatgatga 312

Qy    621 cgaggaggaaggattctgaggatgaagaagaggaggag-ctcttaagaagcctgagccaa 679
          |||||||| |||||||||||||||||||||||||||| ||| ||| ||||||||||||
Db    313 cgaggagg-aggattctgaggatgaagaagaggaggagactcctaataagcctgagccaa 371

Qy    680 tcaacaagaagaggccaaatgaatctgtatccaaaacacccgtctctggaaagaaggcaa 739
          ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db    372 tcaacaagaagaggccaaatgaatctgtatccaraacacccgtctctqgaaagaaggcaa 431
```

PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 34.7%; Score 422.8; DB 21; Length 707;
Best Local Similarity 98.4%; Pred. No. 5.4e-86;
Matches 427; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy    747 agcagcaccagcttctactcctcagaagacagaagagaagaagaaaggaggacacaccgc 806
          ||| | | | ||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    274 agctgaagctgcttctactcctcagaagacagaagagaagaagaaaggaggacacacagc 333

Qy    807 cacaccacacccagctaagaagggtggaaagtctcctgtgaatgctaaccagagccccaa 866
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    334 gacaccacacccagctaagaagggtggaaagtctcctgtgaatgctaaccagagccccaa 393

Qy    867 gtctggaggtcaatcatccggtggtaacaacaacaagaagccattcaactcaggcaaaca 926
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    394 gtctggaggtcaatcatccggtggtaacaacaacaagaagccattcaactcaggcaaaca 453

Qy    927 atttggtggttccaacaacaagggttctaacaagggcaagggaaagggtagagcttaagg 986
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    454 atttggtggttccaacaacaagggttctaacaagggcaagggcaagggtagagcttaagg 513

Qy    987 acgtggatcaaggagaggttttgggttttcgagtagatgatgaaaacacttggaagtgtg 1046
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    514 acgtggatcaaggagaggttttgggttttcgagtagatgatgaaaacacttggaagtgtg 573

Qy   1047 gttttggatttttatcttattttattagtataacttgttatcggatgagctattttgagt 1106
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    574 gttttggatttttatcttattttattagtataacttgttatcggatgagctattttgagt 633

Qy   1107 atttgcaatttctactttcctatgtaattcagtatatgaatatttgctgaaatgagaaag 1166
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    634 atttgcaatttctactttcctatgtaattcagtatatgaatatttgctgaaatgagaaag 693

Qy   1167 aagactcgaattgc 1180
          ||||||||||||||
Db    694 aagactcgaattgc 707
```

RESULT 4
AAC37832
ID AAC37832 standard; DNA; 468 BP.
XX
AC AAC37832;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 18825.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX

PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.

PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.

```
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match             85.1%;  Score 1036.6;  DB 21;  Length 1147;
  Best Local Similarity   98.4%;  Pred. No. 1.6e-224;
  Matches 1121;  Conservative    0;  Mismatches    9;  Indels    9;  Gaps    7;

Qy       2 tctttcgcttctaaaaaaaaacctaacaacctctcttctctcttcctcgttcaacaacaa 61
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 tctttcgcttctaaaaaaaaacctaacaacctctcttctctcttcctcgttcaacaacaa 74

Qy      62 tggagttctggggagttgcggtgacaccaaaaaacgctactaaggtgactcctgaagaag 121
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75 tggagttctggggagttgcggtgacaccaaaaaacgctactaaggtgactcctgaagaag 134

Qy     122 acagccttgtccacatttctcaggcttcacttgactgcacagtgaaatctggagaatctg 181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     135 acagccttgtccacatttctcaggcttcacttgactgcacagtgaaatctggagaatctg 194

Qy     182 tggttttgagtgtgactgttggtggggctaaacttgttattggaacactttcacaagaca 241
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     195 tggttttgagtgtgactgttggtggggctaaacttgttattggaacactttcacaagaca 254

Qy     242 agttccctcagattagctttgatttggtttttgataaagagtttgagctttcacacagcg 301
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     255 agttccctcagattagctttgatttggtttttgataaagagtttgagctttcacacagcg 314

Qy     302 gtaccaaagcaaatgttcatttcattggctacaaatcccccaacatcgagcaggatgact 361
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     315 gtaccaaagcaaatgttcatttcattggctacaaatcccccaacatcgagcaggatgact 374

Qy     362 tcactagttcggatgatgaggatgttcctgaagctgttcctgctcctgcccctactgctg 421
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     375 tcactagttcggatgatgaggatgttcctgaagctgttcctgctcctgcccctactgctg 434

Qy     422 ttactgccaacggaaatgctggagcagctgttgtcaaggctgacacaaagccaaaggcca 481
           |||||||||||||||||||||||||||||||||||||||| ||||||| |||||||
Db     435 ttactgccaacggaaatgctggagcagctgttgtcaaggctgactcaaagcccaaggcca 494

Qy     482 aacctgccgaagtgaagcctgcagaagagaagcctgaatcagacgaggaagatgagtctg 541
           | |||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     495 agcctgccgaagtgaagcctgcagaagagaagccggaatcagacgaggaagatgagtctg 554

Qy     542 atgatgaagatgaagtctgaagaggatgatgactcttgagaaaggaatggatggttgatg 601
           |||||||||||| |||||||||||||||||||| |||||||||||||| |||||||
Db     555 atgatgaagatg-agtctgaagaggatgatgactc-tgagaaaggaatggat-gttgatg 611

Qy     602 aagatgactcagatgatgacgaggaggaaggattctgaggatgaagaagaggaggag-ct 660
           |||||||||||||||||||||||||| |||||||||||||||||||||||||||||| ||
Db     612 aagatgactcagatgatgacgaggagg-aggattctgaggatgaagaagaggaggagact 670

Qy     661 cttaagaagcctgagccaatcaacaagaagaggccaaatgaatctgtatccaaaacaccc 720
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     671 cctaagaagcctgagccaatcaacaagaagaggccaaatgaatctgtatccaaaacaccc 730

Qy     721 gtctctggaaagaaggcaaaaccagcagcagcaccagcttctactcctcagaagacagaa 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     731 gtctctggaaagaaggcaaaaccagcagcagcaccagcttctactcctcagaagacagaa 790

Qy     781 gagaagaagaaaggaggacacaccgccacaccacacccagctaagaagggtggaaagtct 840
           |||||||||||||||||||||| || ||||||||||||||||||||||||||||||||||
Db     791 gagaagaagaaaggaggacacacagcgacaccacacccagctaagaagggtggaaagtct 850

Qy     841 cctgtgaatgctaaccagagccccaagtctggaggtcaatcatccggtggtaacaacaac 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     851 cctgtgaatgctaaccagagccccaagtctggaggtcaatcatccggtggtaacaacaac 910

Qy     901 aagaagccattcaactcaggcaaacaatttggtggttccaacaacaagggttctaacaag 960
           |||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db     911 aagaagccattcaactcaggcaaacaatttggtggttc---caacaagggttctaacaag 967

Qy     961 ggcaagggaaagggtagagcttaaggacgtggatcaaggagaggttttgggttttcgagt 1020
           |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     968 ggcaagggcaagggtagagcttaaggacgtggatcaaggagaggttttgggttttcgagt 1027

Qy    1021 agatgatgaaaacacttggaagtgtggttttggatttttatcttattttattagtataac 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1028 agatgatgaaaacacttggaagtgtggttttggatttttatcttattttattagtataac 1087

Qy    1081 -ttgttatcggatgagctattttgagtatttgcaatttctactttcctatgtaattcag 1138
            |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    1088 tttgttatcggatgagctattttgagtatttgcagtttctactttcctatgtaattcag 1146


RESULT   3
AAC52092
ID   AAC52092 standard; DNA; 707 BP.
XX
AC   AAC52092;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 70362.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
```

PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.

```
DR   WPI; 2001-258457/27.
DR   P-PSDB; AAB67814.
XX
PT   Methods for regulating gene expression in transgenic plants, e.g.
PT   repressing ethylene-responsive phenotypes (e.g. inhibition of hypocotyl
PT   elongation), comprises introducing genes encoding histone deacetylase
PT   -
XX
PS   Claim 9; Fig 2B; 91pp; English.
XX
CC   The present sequence encodes an Arabidopsis thaliana histone deacetylase
CC   designated AtHDA2B. The protein is homologous to yeast RPD3 and HDA1.
CC   The polynucleotide sequence is used in the method of the invention.
CC   The specification describes a method for regulating gene expression in
CC   transgenic plants. The method comprises modifying histones by introducing
CC   chimeric nucleotide sequences which have regulatory elements in operative
CC   association with a gene of interest or with a nucleotide sequence
CC   encoding histone deacetylase. The method is useful for regulating the
CC   developmental, physiological or biochemical pathway within a plant,
CC   particularly for repressing ethylene-responsive phenotypes
CC   (e.g. inhibition of hypocotyl elongation). The method is also useful
CC   as a functional test for identifying a phenotype associated with
CC   perturbing a gene. The histone deacetylase genes are useful for
CC   altering the development of an organism.
XX
SQ   Sequence 1212 BP; 393 A; 238 C; 293 G; 288 T; 0 other;


  Query Match             93.6%;  Score 1139.8;  DB 22;  Length 1212;
  Best Local Similarity   99.2%;  Pred. No. 8e-248;
  Matches 1209;  Conservative    0;  Mismatches    2;  Indels    8;  Gaps    6;

Qy      1 gtctttcgcttctaaaaaaaaacctaacaacctctcttctctcttcctcgttcaacaaca 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 gtctttcgcttctaaaaaaaaacctaacaacctctcttctctcttcctcgttcaacaaca 60

Qy     61 atggagttctggggagttgcggtgacaccaaaaaacgctactaaggtgactcctgaagaa 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 atggagttctggggagttgcggtgacaccaaaaaacgctactaaggtgactcctgaagaa 120

Qy    121 gacagccttgtccacatttctcaggcttcacttgactgcacagtgaaatctggagaatct 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 gacagccttgtccacatttctcaggcttcacttgactgcacagtgaaatctggagaatct 180

Qy    181 gtggttttgagtgtgactgttggtggggctaaacttgttattggaacactttcacaagac 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 gtggttttgagtgtgactgttggtggggctaaacttgttattggaacactttcacaagac 240

Qy    241 aagttccctcagattagctttgatttggtttttgataaagagtttgagctttcacacagc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 aagttccctcagattagctttgatttggtttttgataaagagtttgagctttcacacagc 300

Qy    301 ggtaccaaagcaaatgttcatttcattggctacaaatcccccaacatcgagcaggatgac 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ggtaccaaagcaaatgttcatttcattggctacaaatcccccaacatcgagcaggatgac 360

Qy    361 ttcactagttcggatgatgaggatgttcctgaagctgttcctgctcctgcccctactgct 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ttcactagttcggatgatgaggatgttcctgaagctgttcctgctcctgcccctactgct 420

Qy    421 gttactgccaacggaaatgctggagcagctgttgtcaaggctgacacaaagccaaaggcc 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 gttactgccaacggaaatgctggagcagctgttgtcaaggctgacacaaagccaaaggcc 480

Qy    481 aaacctgccgaagtgaagcctgcagaagagaagcctgaatcagacgaggaagatgagtct 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 aaacctgccgaagtgaagcctgcagaagagaagcctgaatcagacgaggaagatgagtct 540

Qy    541 gatgatgaagatgaagtctgaagaggatgatgactcttgagaaaggaatggatggttgat 600
          |||||||||||| |||||||||||||||||||| ||||||||||||||| ||||||
Db    541 gatgatgaagatg-agtctgaagaggatgatgactc-tgagaaaggaatggat-gttgat 597

Qy    601 gaagatgactcagatgatgacgaggaggaaggattctgaggatgaagaagaggaggag-c 659
          |||||||||||||||||||||||||||||||| |||||||||||||||||||| |||||| |
Db    598 gaagatgactcagatgatgacgaggagg-aggattctgaggatgaagaagaagaggagac 656

Qy    660 tcttaagaagcctgagccaatcaacaagaagaggccaaatgaatctgtatccaaaacacc 719
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    657 tcctaagaagcctgagccaatcaacaagaagaggccaaatgaatctgtatccaaaacacc 716

Qy    720 cgtctctggaaagaaggcaaaaccagcagcagcaccagcttctactcctcagaagacaga 779
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    717 cgtctctggaaagaaggcaaaaccagcagcagcaccagcttctactcctcagaagac--- 773

Qy    780 agagaagaagaaaggaggacacaccgccacaccacacccagctaagaagggtggaaagtc 839
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    774 agagaagaagaaaggaggacacaccgccacaccacacccagctaagaagggtggaaagtc 833

Qy    840 tcctgtgaatgctaaccagagccccaagtctggaggtcaatcatccggtggtaacaacaa 899
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    834 tcctgtgaatgctaaccagagccccaagtctggaggtcaatcatccggtggtaacaacaa 893

Qy    900 caagaagccattcaactcaggcaaacaatttggtggttccaacaacaagggttctaacaa 959
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    894 caagaagccattcaactcaggcaaacaatttggtggttccaacaacaagggttctaacaa 953

Qy    960 gggcaagggaaagggtagagcttaaggacgtggatcaaggagaggttttgggttttcgag 1019
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    954 gggcaagggaaagggtagagcttaaggacgtggatcaaggagaggttttgggttttcgag 1013

Qy   1020 tagatgatgaaaacacttggaagtgtggttttggatttttatcttattttattagtataa 1079
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1014 tagatgatgaaaacactggaagtgtggttttggatttttatcttattttattagtataa 1073

Qy   1080 cttgttatcggatgagctattttgagtatttgcaatttctacttcctatgtaattcagt 1139
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1074 cttgttatcggatgagctattttgagtatttgcaatttctacttcctatgtaattcagt 1133

Qy   1140 atatgaatatttgctgaaatgagaaagaagactcgaattgcaaacaaaaaaaaaaaaaaa 1199
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1134 atatgaatatttgctgaaatgagaaagaagactcgaattgcaaacaaaaaaaaaaaaaa 1193

Qy   1200 aaaaaaaaagggcggccgc 1218
          ||||||||||||||||||
Db   1194 aaaaaaaaagggcggccgc 1212


RESULT   2
AAC41072
ID   AAC41072 standard; DNA; 1147 BP.
XX
AC   AAC41072;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 30546.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 20:26:19 ; Search time 665.94 Seconds
(without alignments)
3140.226 Million cell updates/sec

Title: US-09-645-337-7
Perfect score: 1218
Sequence: 1 gtctttcgcttctaaaaaaa.....aaaaaaaaaagggcggccgc 1218

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1980.DAT:*
2: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1981.DAT:*
3: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1982.DAT:*
4: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1983.DAT:*
5: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1984.DAT:*
6: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1985.DAT:*
7: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1986.DAT:*
8: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1987.DAT:*
9: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1988.DAT:*
10: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1989.DAT:*
11: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1990.DAT:*
12: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1991.DAT:*
13: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1992.DAT:*
14: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1993.DAT:*
15: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1994.DAT:*
16: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1995.DAT:*
17: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:*
18: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:*
19: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:*
20: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:*
21: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:*
22: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:*
23: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:*
24: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1139.8	93.6	1212	22	AAF80353	Nucleotide sequenc
2	1036.6	85.1	1147	21	AAC41072	Arabidopsis thalia
3	422.8	34.7	707	21	AAC52092	Arabidopsis thalia
4	369.4	30.3	468	21	AAC37832	Arabidopsis thalia
5	165.2	13.6	940	21	AAC36931	Arabidopsis thalia
6	162.6	13.3	1002	21	AAC35217	Arabidopsis thalia
7	159.2	13.1	991	21	AAC45610	Arabidopsis thalia
8	156.2	12.8	939	22	AAF80352	Nucleotide sequenc
9	149.8	12.3	774	21	AAC43043	Arabidopsis thalia
10	112	9.2	1307	20	AAX90845	Maize histone deac
11	111.6	9.2	1191	20	AAX90843	Maize histone deac
12	111.2	9.1	1245	20	AAX90844	Maize histone deac
13	111.2	9.1	1283	20	AAX90842	Maize histone deac
14	110.4	9.1	331	21	AAC52921	Arabidopsis thalia
15	90	7.4	741	21	AAC40930	Arabidopsis thalia
16	66.2	5.4	856	21	AAC49668	Arabidopsis thalia
17	66.2	5.4	879	21	AAC34537	Arabidopsis thalia
18	59.8	4.9	690	23	AAS90703	DNA encoding novel
19	59.8	4.9	693	23	AAS74240	DNA encoding novel
20	59.8	4.9	693	23	AAS90715	DNA encoding novel
21	58.4	4.8	282	23	AAS90681	DNA encoding novel
c 22	57.8	4.7	360	22	ABA71815	Human foetal liver
c 23	57.8	4.7	360	22	AAK20192	Human brain expres
c 24	57.8	4.7	360	22	AAK46257	Human bone marrow
c 25	57.8	4.7	360	22	AAI52158	Probe #20844 used
26	57.8	4.7	1276	21	AAA15086	Transcript SCP-3B
27	56.8	4.7	425	22	AAH33149	Human colon cancer
28	56.6	4.6	1686	16	AAQ87587	DNA encoding Leuco
c 29	55.8	4.6	450	22	ABA59302	Human foetal liver
c 30	55.8	4.6	450	22	AAK07521	Human brain expres
c 31	55.8	4.6	450	22	AAK33313	Human bone marrow
c 32	55.8	4.6	450	22	AAI39102	Probe #7788 used t
33	55.4	4.5	932	21	AAC59530	Human secreted pro
c 34	54.6	4.5	612	22	AAH71471	Human cervical can
35	54	4.4	2814	23	AAS65681	DNA encoding novel
c 36	53.6	4.4	411	22	AAI88595	Human polynucleoti
c 37	53.4	4.4	337	22	AAL11578	Human breast cance
c 38	52.8	4.3	351	22	AAL14805	Human breast cance
39	52.8	4.3	966	20	AAX60314	DNA encoding a hum
40	52.6	4.3	198	22	AAS07727	Cervical cancer pr
41	52.4	4.3	393	23	AAS90684	DNA encoding novel
42	52.2	4.3	318	21	AAC10202	Human secreted pro
43	52	4.3	2334	23	AAS90705	DNA encoding novel
44	51.8	4.3	447	23	AAS75731	DNA encoding novel
45	51.8	4.3	510	23	AAS69539	DNA encoding novel

ALIGNMENTS

RESULT 1
AAF80353
ID AAF80353 standard; DNA; 1212 BP.
XX
AC AAF80353;
XX
DT 29-JUN-2001 (first entry)
XX
DE Nucleotide sequence of a histone deacetylase designated AtHDA2B.
XX
KW Histone deacetylase; AtHDA2B; RPD3; gene expression; transgenic plant;
KW HDA1; ethylene-responsive phenotype; hypocotyl elongation; ds.
XX
OS Arabidopsis thaliana.
XX
FH Key Location/Qualifiers
FT CDS 61..978
FT /*tag= a
FT /product= "histone deacetylase AtHDA2B"
XX
PN CA2316036-A1.
XX
PD 27-FEB-2001.
XX
PF 24-AUG-2000; 2000CA-2316036.
XX
PR 27-AUG-1999; 99US-0383971.
XX
PA (MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.
XX
PI Miki B, Brown D, Tian L, Wu K;
XX

```
RESULT 10
BE523968
LOCUS       BE523968                422 bp    mRNA    linear   EST 19-MAR-2001
DEFINITION  M44A9STM Arabidopsis developing seed Arabidopsis thaliana cDNA
            clone M44A9 5', mRNA sequence.
ACCESSION   BE523968
VERSION     BE523968.1  GI:9781946
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 422)
  AUTHORS   White,J.A., Todd,J., Newman,T., Focks,N., Girke,T., Martinez de
            Ilarduya,O., Jaworski,J.G., Ohlrogge,J. and Benning,C.
  TITLE     A new set of Arabidopsis expressed sequence tags from developing
            seeds. The metabolic pathway from carbohydrates to seed oil
  JOURNAL   Plant Physiol. 124 (4), 1582-1594 (2000)
  MEDLINE   20567808
COMMENT     Contact: Benning, C
            Dept. of Biochemistry & Molecular Biology
            Michigan State University
            224 Biochemistry, Michigan State University, East Lansing, MI 48824
            , USA
            Tel: 517 355 1609
            Fax: 517 353 9334
            Email: benning@msu.edu
            Michigan State University DNA Sequencing Facility Arabidopsis
            Biological Resource Center, The Ohio State University, 309 Botany &
            Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210 USA, FAX:
            6142920603 TEL: 6142929371.
FEATURES             Location/Qualifiers
     source          1. .422
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /clone="M44A9"
                     /clone_lib="Arabidopsis developing seed"
                     /tissue_type="seed"
                     /dev_stage="5-13 days after flowering"
                     /lab_host="E.coli"
                     /note="Organ: Developing seed; Vector: pBluescript SK-;
                     Site_1: EcoRI; Site_2: XhoII"
BASE COUNT      109 a     99 c     92 g    122 t
ORIGIN
```

```
  Query Match           33.9%;  Score 412.4;  DB 10;  Length 422;
  Best Local Similarity  99.8%;  Pred. No. 2.6e-48;
  Matches  413;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      15 aaaaaaaacctaacaacctctcttctctcttcctcgttcaacaacaatggagttctgggg 74
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9 AGAAAAAACCTAACAACCTCTCTTCTCTCTTCCTCGTTCAACAACAATGGAGTTCTGGGG 68

Qy      75 agttgcggtgacaccaaaaaacgctactaaggtgactcctgaagaagacagccttgtcca 134
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69 AGTTGCGGTGACACCAAAAAACGCTACTAAGGTGACTCCTGAAGAAGACAGCCTTGTCCA 128

Qy     135 catttctcaggcttcacttgactgcacagtgaaatctggagaatctgtggttttgagtgt 194
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129 CATTTCTCAGGCTTCACTTGACTGCACAGTGAAATCTGGAGAATCTGTGGTTTTGAGTGT 188

Qy     195 gactgttggtggggctaaacttgttattggaacactttcacaagacaagttcccctcagat 254
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     189 GACTGTTGGTGGGGCTAAACTTGTTATTGGAACACTTTCACAAGACAAGTTCCCTCAGAT 248

Qy     255 tagctttgatttggtttttgataaagagtttgagctttcacacagcggtaccaaaagcaaa 314
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     249 TAGCTTTGATTTGGTTTTTGATAAAGAGTTTGAGCTTTCACACAGCGGTACCAAAGCAAA 308

Qy     315 tgttcatttcattggctacaaatcccccaacatcgagcaggatgacttcactagttcgga 374
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     309 TGTTCATTTCATTGGCTACAAATCCCCCAACATCGAGCAGGATGACTTCACTAGTTCGGA 368

Qy     375 tgatgaggatgttcctgaagctgttcctgctcctgcccctactgctgttactgc 428
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     369 TGATGAGGATGTTCCTGAAGCTGTTCCTGCTCCTGCCCCTACTGCTGTTACTGC 422
```

```
RESULT 11
AV542939/c
LOCUS       AV542939                421 bp    mRNA    linear   EST 06-SEP-2000
DEFINITION  AV542939 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
            cDNA clone RZ194a06F 3', mRNA sequence.
ACCESSION   AV542939
VERSION     AV542939.1  GI:8704699
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 421)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     A large scale analysis of cDNA in Arabidopsis thaliana: Generation
            of 12,028 non-redundant expressed sequence tags from normalized and
            size-selected cDNA libraries
  JOURNAL   DNA Res. 7, 175-180 (2000)
  MEDLINE   20363093
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .421
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /clone="RZ194a06F"
                     /clone_lib="Arabidopsis thaliana roots Columbia"
                     /tissue_type="roots"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT      111 a    102 c     68 g    140 t
ORIGIN
```

```
  Query Match           33.8%;  Score 411.4;  DB 9;  Length 421;
  Best Local Similarity  99.8%;  Pred. No. 3.6e-48;
  Matches  412;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     777 agaagagaagaagaaaggaggacacaccgccacaccacacccagctaagaagggtggaaa 836
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 AGAAGAGAAGAAGAAAGGAGGACACACCGCCACACCACACCCAGCTAAGAAGGGTGGAAA 362

Qy     837 gtctcctgtgaatgctaaccagagccccaagtctggaggtcaatcatccggtggtaacaa 896
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GTCTCCTGTGAATGCTAACCAGAGCCCCAAGTCTGGAGGTCAATCATCCGGTGGTAACAA 302

Qy     897 caacaagaagccattcaactcaggcaaacaatttggtggttccaacaacaagggttctaa 956
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CAACAAGAAGCCATTCAACTCAGGCAAACAATTTGGTGGTTCCAACAACAAGGGTTCTAA 242

Qy     957 caagggcaagggaaagggtagagcttaaggacgtggatcaaggagaggttttgggttttc 1016
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CAAGGGCAAGGGAAAGGGTAGAGCTTAAGGACGTGGATCAAGGAGAGGTTTTGGGTTTTC 182

Qy    1017 gagtagatgatgaaaacacttggaagtgtggttttggatttttatcttattttattagta 1076
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GAGTAGATGATGAAAACACTTGGAAGTGTGGTTTTGGATTTTTATCTTATTTTATTAGTA 122

Qy    1077 taacttgttatcggatgagctattttgagtatttgcaatttctactttcctatgtaattc 1136
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    121 TAACTTGTTATCGGATGAGCTATTTTGAGTATTTGCAATTTCTACTTTCCTATGTAATTC 62

Qy   1137 agtatatgaatatttgctgaaatgagaaagaagactcgaattgcaaacaaaaa 1189
          ||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db     61 AGTATATGAATATTTGCTGAAATGAGAAAGAAGACTCGAATTGCAAACACAAA 9


RESULT  12
T45874
LOCUS       T45874                  504 bp    mRNA    linear   EST 09-JAN-1998
DEFINITION  9137 Lambda-PRL2 Arabidopsis thaliana cDNA clone 133F15T7, mRNA
            sequence.
ACCESSION   T45874
VERSION     T45874.1  GI:934127
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 504)
  AUTHORS   Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh
            ,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M., Retzel
            ,E. and Somerville,C.
  TITLE     Genes galore: a summary of methods for accessing results from
            large-scale partial sequencing of anonymous Arabidopsis cDNA clones
  JOURNAL   Plant Physiol. 106, 1241-1255 (1994)
  MEDLINE   95148729
COMMENT     On Apr 14, 1993 this sequence version replaced gi:638462.
            Contact: Thomas Newman
            MSU-DOE Plant Research Laboratory
            Michigan State University
            MSU-DOE-PRL, Michigan State University,Plant Biology Bldg.,E.
            Lansing,MI
            Tel: 517-353-0854
            Fax: 517-353-9168
            Email: 22313tcn@ibm.cl.msu.edu
            Seq primer: T7 dye primer.
FEATURES             Location/Qualifiers
     source          1. .504
                     /organism="Arabidopsis thaliana"
                     /strain="var columbia"
                     /db_xref="taxon:3702"
                     /clone="133F15T7"
                     /clone_lib="Lambda-PRL2"
                     /note="Vector: lambda Zip-Lox; Site_1: Sal; Site_2: Not;
                     Lambda PRL2 is a cDNA library derived from equal
                     quantities of 4 pools of mRNA.  The mRNA sources were 1) 7
                     day germinated etiolated seedlings; 2) tissue culture
                     grown roots; 3) staged plants half with 24 hour light
                     cycle, half on 16 hr light, 8 hour dark- rosettes;  4)
                     same plants as 3 but aerial tissue (stems, flowers and
                     siliques.  The vector is BRL's lambda Zip-Lox.  The cDNA
                     inserts were directionally cloned with Sal-Not arms using
                     oligo dT primed cDNA.  "
BASE COUNT      131 a    113 c    109 g    142 t      9 others
ORIGIN


  Query Match           33.6%;  Score 409.4;  DB 10;  Length 504;
  Best Local Similarity   95.9%;  Pred. No. 6.4e-48;
  Matches  439;  Conservative    0;  Mismatches   16;  Indels    3;  Gaps    2;

Qy       2 tctttcgcttctaaaaaaaaacctaacaacctctcttctctcttcctcgttcaacaacaa 61
           |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db       1 TCTTTCGCTTCTAAAAAAAAACCTANCAACCTCTCTTCTCTCTTCCTCGTTCAACAACAA 60

Qy      62 tggagttctggggagttgcggtgacaccaaaaaacgctactaaggtgactcctgaagaag 121
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TGGAGTTCTGGGGAGTTGCGGTGACACCAAAAAACGCTACTAAGGTGACTCCTGAAGAAG 120

Qy     122 acagccttgtccacatttctcaggcttcacttgactgcacagtgaaatctggagaatctg 181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db     121 ACAGCCTTGTCCACATTTCTCAGGCTTCACTTGACTGCACAGTGAAATCTGGAGAATCTG 180

Qy     182 tggttttgagtgtgactgttggtggggctaaacttgttattggaacactttcacaagaca 241
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TGGTTTTGAGTGTGACTGTTGGTGGGGCTAAACTTGTTATTGGAACACTTTCACAAGACA 240

Qy     242 agttccctcagattagctttgatttggtttttgataaagagtttgagctttcacacagcg 301
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AGTTCCCTCAGATTAGCTTTGATTTGGTTTTTGATAAAGAGTTTGAGCTTTCACACAGCG 300

Qy     302 gtaccaaagcaaatgttcatttcattggctacaaatcccccaacatcgagcaggatgact 361
           |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db     301 GTACCAAAGCAAATNTTCATTTCATTGGCTACAAATCCCCCAACATCGAGCAGGATGACT 360

Qy     362 tcactagttcggatgatgaggatgttcctgaagctgttcctgctcctgccccctactgctg 421
           |||||||||||||||||||||||||||||||||||||||||| ||   |||  |||| ||
Db     361 TCACTAGTTCGGATGATGAGGATGTTCCTGAAGCTGTTCCTGTTCTGNCCC--ACTGNTG 418

Qy     422 ttactgccaacggaaatgctggagcagctgttgtcaag 459
           |||||||| |||| ||| |||| |||| | ||   | |
Db     419 TTACTGCC-ACGGGAATCCTGGGGCAGTTTTTNCAAGG 455


RESULT  13
AA597939
LOCUS       AA597939                444 bp    mRNA    linear   EST 19-SEP-1997
DEFINITION  29188 Lambda-PRL2 Arabidopsis thaliana cDNA clone 248C11T7, mRNA
            sequence.
ACCESSION   AA597939
VERSION     AA597939.1  GI:2413362
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 444)
  AUTHORS   Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh
            ,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M., Retzel
            ,E. and Somerville,C.
  TITLE     Genes galore: a summary of methods for accessing results from
            large-scale partial sequencing of anonymous Arabidopsis cDNA clones
  JOURNAL   Plant Physiol. 106, 1241-1255 (1994)
  MEDLINE   95148729
COMMENT     Contact: Thomas Newman
            MSU-DOE Plant Research Laboratory
            Michigan State University
            MSU-DOE-PRL, Michigan State University,Plant Biology Bldg.,E.
            Lansing,MI
            Tel: 517-353-0854
            Fax: 517-353-9168
            Email: 22313tcn@ibm.cl.msu.edu
            Seq primer: T7 dye primer.
FEATURES             Location/Qualifiers
     source          1. .444
                     /organism="Arabidopsis thaliana"
                     /strain="var columbia"
                     /db_xref="taxon:3702"
                     /clone="248C11T7"
                     /clone_lib="Lambda-PRL2"
                     /note="Vector: lambda Zip-Lox; Site_1: Sal; Site_2: Not;
                     Lambda PRL2 is a cDNA library derived from equal
                     quantities of 4 pools of mRNA.  The mRNA sources were 1) 7
                     day germinated etiolated seedlings; 2) tissue culture
                     grown roots; 3) staged plants half with 24 hour light
                     cycle, half on 16 hr light, 8 hour dark- rosettes;  4)
                     same plants as 3 but aerial tissue (stems, flowers and
                     siliques.  The vector is BRL's lambda Zip-Lox.  The cDNA
                     inserts were directionally cloned with Sal-Not arms using
                     oligo dT primed cDNA.  "
BASE COUNT      138 a     84 c    109 g    106 t      7 others
ORIGIN
```

```
  Query Match             33.4%;  Score 407;  DB 9;  Length 444;
  Best Local Similarity   95.9%;  Pred. No. 1.5e-47;
  Matches  424;  Conservative    0;  Mismatches   17;  Indels    1;  Gaps    1;

Qy    741 accagcagcagcaccagcttctactcctcagaagacagaagagaagaagaaagaaggaggaca 800
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db      1 ACCAGCAGCAGCACCAGCTTCTACNCCTCAGAAGACAGAAGAGAAGAAGAAAGGAGGACA 60

Qy    801 caccgccacaccacacccagctaagaagggtggaaagtctcctgtgaatgctaaccagag 860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CACCGCCACACCACACCCAGCTAAGAAGGGTGGAAAGTCTCCTGTGAATGCTAACCAGAG 120

Qy    861 ccccaagtctggaggtcaatcatccggtggtaacaacaacaagaagccattcaactcagg 920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCCCAAGTCTGGAGGTCAATCATCCGGTGGTAACAACAACAAGAAGCCATTCAACTCAGG 180

Qy    921 caaacaatttggtggttccaacaacaagggttctaacaagggcaagggaaagggtagagc 980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CAAACAATTTGGTGGTTCCAACAACAAGGGTTCTAACAAGGGCAAGGGAAAGGGTAGAGC 240

Qy    981 ttaaggacgtggatcaaggagaggttttgggttttcgagtagatgatgaaaacacttgga 1040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TTAAGGACGTGGATCAAGGAGAGGTTTTGGGTTTTCGAGTAGATGATGAAAACACTTGGA 300

Qy   1041 agtgtggttttggatttttatcttattttattagtataacttgttatc-ggatgagctat 1099
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    301 AGTGTGGTTTTGGATTTTTATCTTATTTTATTAGTATAACTTGTTATCGGGATGAGCTAT 360

Qy   1100 tttgagtatttgcaatttctactttcctatgtaattcagtatatgaatatttgctgaaat 1159
          |||||||||||||||||| ||||||| ||| | | ||| ||||| |||||||| ||| |
Db    361 TTTGAGTATTTGCAATTTCAACTTTCCCATGNATTCCAGNATATGGATATTTGCCGAATT 420

Qy   1160 gagaaagaagactcgaattgca 1181
          | | ||||||   ||| || ||
Db    421 GGGGAAGAAGNNNCGATTTNCA 442


RESULT  14
AV537871/c
LOCUS       AV537871                422 bp    mRNA    linear   EST 06-SEP-2000
DEFINITION  AV537871 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
            cDNA clone RZ106a10F 3', mRNA sequence.
ACCESSION   AV537871
VERSION     AV537871.1  GI:8698154
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 422)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     A large scale analysis of cDNA in Arabidopsis thaliana: Generation
            of 12,028 non-redundant expressed sequence tags from normalized and
            size-selected cDNA libraries
  JOURNAL   DNA Res. 7, 175-180 (2000)
  MEDLINE   20363093
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .422
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /clone="RZ106a10F"
                     /clone_lib="Arabidopsis thaliana roots Columbia"
                     /tissue_type="roots"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT      114 a    100 c     69 g    139 t
ORIGIN


  Query Match             33.4%;  Score 406.4;  DB 9;  Length 422;
  Best Local Similarity   99.8%;  Pred. No. 1.8e-47;
  Matches  407;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    782 agaagaagaaaggaggacacaccgccacaccacacccagctaagaagggtggaaagtctc 841
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    422 AGAAGAAGAAAGGAGGACACACCGCCACACCACACCCAGCTAAGAAGGGTGGAAAGTCTC 363

Qy    842 ctgtgaatgctaaccagagccccaagtctggaggtcaatcatccggtggtaacaacaaca 901
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    362 CTGTGAATGCTAACCAGAGCCCCAAGTCTGGAGGTCAATCATCCGGTGGTAACAACAACA 303

Qy    902 agaagccattcaactcaggcaaacaatttggtggttccaacaacaagggttctaacaagg 961
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    302 AGAAGCCATTCAACTCAGGCAAACAATTTGGTGGTTCCAACAACAAGGGTTCTAACAAGG 243

Qy    962 gcaagggaaagggtagagcttaaggacgtggatcaaggagaggttttgggttttcgagta 1021
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    242 GCAAGGGAAAGGGTAGAGCTTAAGGACGTGGATCAAGGAGAGGTTTTGGGTTTTCGAGTA 183

Qy   1022 gatgatgaaaacacttggaagtgtggttttggatttttatcttatttttattagtataact 1081
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    182 GATGATGAAAACACTTGGAAGTGTGGTTTTGGATTTTTATCTTATTTTTATTAGTATAACT 123

Qy   1082 tgttatcggatgagctattttgagtatttgcaatttctactttcctatgtaattcagtat 1141
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    122 TGTTATCGGATGAGCTATTTTGAGTATTTGCAATTTCTACTTTCCTATGTAATTCAGTAT 63

Qy   1142 atgaatatttgctgaaatgagaaagaagactcgaattgcaaacaaaaa 1189
          ||||||||||||||||||||||||||||||||||||||||||||||| |||
Db     62 ATGAATATTTGCTGAAATGAGAAAGAAGACTCGAATTGCAAACACAAA 15


RESULT  15
AV541169/c
LOCUS       AV541169                407 bp    mRNA    linear   EST 07-SEP-2000
DEFINITION  AV541169 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
            cDNA clone RZ160c03F 3', mRNA sequence.
ACCESSION   AV541169
VERSION     AV541169.1  GI:8702927
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 407)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     A large scale analysis of cDNA in Arabidopsis thaliana: Generation
            of 12,028 non-redundant expressed sequence tags from normalized and
            size-selected cDNA libraries
  JOURNAL   DNA Res. 7, 175-180 (2000)
  MEDLINE   20363093
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .407
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /clone="RZ160c03F"
                     /clone_lib="Arabidopsis thaliana roots Columbia"
                     /tissue_type="roots"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT      106 a    100 c     67 g    131 t      3 others
```

ORIGIN

Query Match 33.2%; Score 404; DB 9; Length 407;
Best Local Similarity 99.3%; Pred. No. 3.9e-47;
Matches 404; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy    778 gaagagaagaagaaaggaggacacaccgccacaccacacccagctaagaagggtggaaag 837
          |||||||||||||||||||| | | ||||||||||||||||||||||||||||||||||||
Db    407 GAAGAGAAGAAGAAAGGAGGNCNCNCCGCCACACCACACCCAGCTAAGAAGGGTGGAAAG 348

Qy    838 tctcctgtgaatgctaaccagagccccaagtctggaggtcaatcatccggtggtaacaac 897
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    347 TCTCCTGTGAATGCTAACCAGAGCCCCAAGTCTGGAGGTCAATCATCCGGTGGTAACAAC 288

Qy    898 aacaagaagccattcaactcaggcaaacaatttggtggttccaacaacaagggttctaac 957
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    287 AACAAGAAGCCATTCAACTCAGGCAAACAATTTGGTGGTTCCAACAACAAGGGTTCTAAC 228

Qy    958 aagggcaagggaaagggtagagcttaaggacgtggatcaaggagaggttttgggttttcg 1017
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    227 AAGGGCAAGGGAAAGGGTAGAGCTTAAGGACGTGGATCAAGGAGAGGTTTTGGGTTTTCG 168

Qy   1018 agtagatgatgaaaacacttggaagtgtggttttggatttttatcttatttttattagtat 1077
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    167 AGTAGATGATGAAAACACTTGGAAGTGTGGTTTTGGATTTTTATCTTATTTTATTAGTAT 108

Qy   1078 aacttgttatcggatgagctattttgagtatttgcaatttctactttcctatgtaattca 1137
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107 AACTTGTTATCGGATGAGCTATTTTGAGTATTTGCAATTTCTACTTTCCTATGTAATTCA 48

Qy   1138 gtatatgaatatttgctgaaatgagaaagaagactcgaattgcaaac 1184
          |||||||||||||||||||||||||||||||||||||||||||||
Db     47 GTATATGAATATTTGCTGAAATGAGAAAGAAGACTCGAATTGCAAAC 1
```

Search completed: May 2, 2002, 18:43:53
Job time: 10877 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2002, 22:03:52 ; Search time 144.83 Seconds
(without alignments)
233.912 Million cell updates/sec

Title: US-09-645-337-8
Perfect score: 1589
Sequence: 1 MEFWGVAVTPKNATKVTPEE.....KQFGGSNNKGSNKGKGKGRA 305

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1589	100.0	305	22	AAB67814	Amino acid sequenc
2	1558	98.0	305	21	AAG26185	Arabidopsis thalia
3	740	46.6	155	21	AAG17704	Arabidopsis thalia
4	594	37.4	245	21	AAG38221	Arabidopsis thalia
5	594	37.4	245	22	AAB67813	Amino acid sequenc
6	594	37.4	257	21	AAG31484	Arabidopsis thalia
7	589.5	37.1	248	21	AAG10894	Arabidopsis thalia
8	544	34.2	311	20	AAY28804	Maize histone deac
9	535	33.7	302	20	AAY28803	Maize histone deac
10	528	33.2	305	20	AAY28802	Maize histone deac
11	436	27.4	285	20	AAY28805	Maize histone deac
12	393.5	24.8	149	21	AAG54981	Arabidopsis thalia
13	345.5	21.7	208	21	AAG25823	Arabidopsis thalia
14	311	19.6	412	20	AAW68010	S. frugiperda immu
15	262.5	16.5	203	21	AAG09171	Arabidopsis thalia
16	227	14.3	349	22	ABB62190	Drosophila melanog
17	223.5	14.1	194	21	AAG54925	Arabidopsis thalia
18	219	13.8	98	21	AAG15354	Arabidopsis thalia
19	216.5	13.6	181	21	AAG09172	Arabidopsis thalia
20	212.5	13.4	126	21	AAG15353	Arabidopsis thalia
21	205	12.9	707	16	AAR79912	Human nucleolin.
22	205	12.9	707	20	AAW84052	Human V3 loop HIV
23	205	12.9	707	22	AAB48964	Human nucleolin.
24	200.5	12.6	570	22	ABB57851	Drosophila melanog
25	194.5	12.2	654	22	ABB63266	Drosophila melanog
26	194.5	12.2	688	22	ABB63269	Drosophila melanog
27	193.5	12.2	147	21	AAG09173	Arabidopsis thalia
28	190.5	12.0	536	22	AAB92836	Human protein sequ
29	189	11.9	857	22	ABB63922	Drosophila melanog
30	188	11.8	723	14	AAR38745	Drosophila SSRP co
31	188	11.8	723	19	AAW39213	D. melanogaster SS
32	188	11.8	723	22	ABB61000	Drosophila melanog
33	184.5	11.6	682	21	AAG46506	Arabidopsis thalia
34	184.5	11.6	724	21	AAG46505	Arabidopsis thalia
35	184.5	11.6	778	21	AAG46504	Arabidopsis thalia
36	180	11.3	1162	21	AAY96255	Kaposi's sarcoma-a
37	180	11.3	1162	21	AAY58500	HHV8 ORF 73 protei
38	180	11.3	1162	22	AAB62331	Amino acid sequenc
39	179.5	11.3	670	21	AAG51795	Arabidopsis thalia
40	179.5	11.3	688	21	AAG51794	Arabidopsis thalia
41	179.5	11.3	722	21	AAG51793	Arabidopsis thalia
42	178.5	11.2	517	22	AAM39867	Human polypeptide
43	178.5	11.2	520	22	AAM41653	Human polypeptide
44	177.5	11.2	885	21	AAB18279	Plasmodium falcipa
45	174.5	11.0	718	21	AAG42569	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAB67814
ID AAB67814 standard; Protein; 305 AA.
XX
AC AAB67814;
XX
DT 29-JUN-2001 (first entry)
XX
DE Amino acid sequence of a histone deacetylase designated AtHDA2B.
XX
KW Histone deacetylase; AtHDA2B; RPD3; gene expression; transgenic plant;
KW HDA1; ethylene-responsive phenotype; hypocotyl elongation.
XX
OS Arabidopsis thaliana.
XX
PN CA2316036-A1.
XX
PD 27-FEB-2001.
XX
PF 24-AUG-2000; 2000CA-2316036.
XX
PR 27-AUG-1999; 99US-0383971.
XX
PA (MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.
XX
PI Miki B, Brown D, Tian L, Wu K;
XX
DR WPI; 2001-258457/27.
DR N-PSDB; AAF80353.
XX
PT Methods for regulating gene expression in transgenic plants, e.g.
PT repressing ethylene-responsive phenotypes (e.g. inhibition of hypocotyl
PT elongation), comprises introducing genes encoding histone deacetylase
PT -

```
XX
PS   Claim 10; Fig 2B; 91pp; English.
XX
CC   The present sequence represents Arabidopsis thaliana histone deacetylase
CC   designated AtHDA2B. The protein is homologous to yeast RPD3 and HDA1.
CC   The polynucleotide sequence is used in the method of the invention.
CC   The specification describes a method for regulating gene expression in
CC   transgenic plants. The method comprises modifying histones by introducing
CC   chimeric nucleotide sequences which have regulatory elements in operative
CC   association with a gene of interest or with a nucleotide sequence
CC   encoding histone deacetylase. The method is useful for regulating the
CC   developmental, physiological or biochemical pathway within a plant,
CC   particularly for repressing ethylene-responsive phenotypes
CC   (e.g. inhibition of hypocotyl elongation). The method is also useful
CC   as a functional test for identifying a phenotype associated with
CC   perturbing a gene. The histone deacetylase genes are useful for
CC   altering the development of an organism.
XX
SQ   Sequence   305 AA;


  Query Match             100.0%;  Score 1589;  DB 22;  Length 305;
  Best Local Similarity   100.0%;  Pred. No. 4.7e-119;
  Matches  305;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQD 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mefwgvavtpknatkvtpeedslvhisqasldctvksgesvvlsvtvggaklvigtlsqd 60

Qy      61 KFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 kfpqisfdlvfdkefelshsgtkanvhfigykspnieqddftssddedvpeavpapapta 120

Qy     121 VTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMDVDE 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 vtangnagaavvkadtkpkakpaevkpaeekpesdeedesddedeseedddsekgmdvde 180

Qy     181 DDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTEK 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ddsdddeeedsedeeeeetpkkpepinkkrpnesvsktpvsgkkakpaaapastpqktek 240

Qy     241 KKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGGNNNKKPFNSGKQFGGSNNKGSNKGK 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 kkgghtatphpakkggkspvnanqspksgggqssggnnnkkpfnsgkqfggsnnkgsnkgk 300

Qy     301 GKGRA 305
           |||||
Db     301 gkgra 305


RESULT   2
AAG26185
ID   AAG26185 standard; Protein; 305 AA.
XX
AC   AAG26185;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 30547.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
```

PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 98.0%; Score 1558; DB 21; Length 305;
Best Local Similarity 99.0%; Pred. No. 1.4e-116;
Matches 303; Conservative 1; Mismatches 0; Indels 2; Gaps 2;

```
Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQD 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mefwgvavtpknatkvtpeedslvhisqasldctvksgesvvlsvtvggaklvigtlsqd 60

Qy      61 KFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 kfpqisfdlvfdkefelshsgtkanvhfigykspnieqddftssddedvpeavpapapta 120

Qy     121 VTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMDVDE 180
           |||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db     121 vtangnagaavvkadskpkakpaevkpaeekpesdeedesddedeseedddsekgmdvde 180

Qy     181 DDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQKT-E 239
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     181 ddsdddeeedsedeeeeetpkkpepinkkrpnesvsktpvsgkkakpaaapastpqktee 240

Qy     240 KKKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGGNNNKKPFNSGKQFGGSNNKGSNKG 299
           ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db     241 kkkgghtatphpakkggkspvnanqspksggqssggnnnkkpfnsgkqfggs-nkgsnkg 299

Qy     300 KGKGRA 305
           ||||||
Db     300 kgkgra 305
```

RESULT 3
AAG17704
ID AAG17704 standard; Protein; 155 AA.
XX
AC AAG17704;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 18826.
XX

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.

```
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match             46.6%;  Score 740;  DB 21;  Length 155;
  Best Local Similarity   95.3%;  Pred. No. 1.3e-51;
  Matches  143;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      88 FIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTAVTANGNAGAAVVKADTKPKAKPAEVKP 147
           |   | ||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db       6 faskkkpnieqddftssddedvpeavpapaptavtangnagaavvkadskpkakpaevkp 65

Qy     148 AEEKPESDEEDESDDEDESEEDDDSEKGMDVDEDDSDDDEEEDSEDEEEEETPKKPEPIN 207
           |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db      66 aeekpesdeedesddedeseedddsekgmdvdeddsdddeeedsedeeeeetpnkpepin 125

Qy     208 KKRPNESVSKTPVSGKKAKPAAAPASTPQK 237
           ||||||||| ||||||||||||||||||||
Db     126 kkrpnesvsxtpvsgkkakpaaapastpqk 155


RESULT   4
AAG38221
ID   AAG38221 standard; Protein; 245 AA.
XX
AC   AAG38221;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 47121.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
```

PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 37.4%; Score 594; DB 21; Length 245;
Best Local Similarity 50.7%; Pred. No. 1e-39;
Matches 151; Conservative 26; Mismatches 63; Indels 58; Gaps 15;

```
Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASL-DCTVKSGESVVLSVTVGGAKLVIGTLSQ 59
           |||||: |       |||||  |:|:|||||| :|  | || | | | ||   ||:|||| 
Db       1 mefwgievksgkpvtvtpeegilihvsqaslgecknkkgefvplhvkvgnqnlvlgtlst 60

Qy      60 DKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPT 119
           :  ||:  |||||||||||||:  | :|:|:|||:||||   : | ::|:  | |||    
Db      61 enipqlfcdlvfdkefelshtwgkgsvyfvgyktpniepqgy-seeeeeeeeevpa---- 115

Qy     120 AVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMDVD 179
               |||  ||     |||||||||||| :    ||||||| :           ||  |  
Db     116 -----gnaakav----akpkakpaevkpavd----deedesdsd----------gm--d 149

Qy     180 EDDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTE 239
           |||||    |||  |||| |||||   :||| ||:  | ||| |||| |     |||||:  
Db     150 eddsdg---eds--eeeeptpkkpas-skkranettpkapvsakkakvav----tpqktd 199

Qy     240 KKKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGGNNNKKPFNSGKQFGGSNNKGSN 297
           :|          |||||:   |||||||  | | | : || ||||     |:||  :    
Db     200 ek----------kkggka---anqspksasqvscg-sckktfnsgnal-eshnkakh 241
```

```
RESULT   5
AAB67813
ID   AAB67813 standard; Protein; 245 AA.
XX
AC   AAB67813;
XX
DT   29-JUN-2001  (first entry)
XX
DE   Amino acid sequence of a histone deacetylase designated AtHDA2A.
XX
KW   Histone deacetylase; AtHDA2A; RPD3; gene expression; transgenic plant;
KW   HDA1; ethylene-responsive phenotype; hypocotyl elongation.
XX
OS   Arabidopsis thaliana.
XX
PN   CA2316036-A1.
XX
PD   27-FEB-2001.
XX
PF   24-AUG-2000; 2000CA-2316036.
XX
PR   27-AUG-1999;   99US-0383971.
XX
PA   (MIAC ) CANADA MIN AGRIC & AGRI-FOOD CANADA.
XX
PI   Miki B,  Brown D,  Tian L,  Wu K;
XX
DR   WPI; 2001-258457/27.
DR   N-PSDB; AAF80352.
XX
PT   Methods for regulating gene expression in transgenic plants, e.g.
PT   repressing ethylene-responsive phenotypes (e.g. inhibition of hypocotyl
PT   elongation), comprises introducing genes encoding histone deacetylase
PT   -
XX
PS   Claim 10; Fig 2A; 91pp; English.
XX
CC   The present sequence represents Arabidopsis thaliana histone deacetylase
CC   designated AtHDA2A. The protein is homologous to yeast RPD3 and HDA1.
CC   The polynucleotide sequence is used in the method of the invention.
CC   The specification describes a method for regulating gene expression in
CC   transgenic plants. The method comprises modifying histones by introducing
CC   chimeric nucleotide sequences which have regulatory elements in operative
CC   association with a gene of interest or with a nucleotide sequence
CC   encoding histone deacetylase. The method is useful for regulating the
CC   developmental, physiological or biochemical pathway within a plant,
CC   particularly for repressing ethylene-responsive phenotypes
CC   (e.g. inhibition of hypocotyl elongation). The method is also useful
CC   as a functional test for identifying a phenotype associated with
CC   perturbing a gene. The histone deacetylase genes are useful for
CC   altering the development of an organism.
XX
SQ   Sequence   245 AA;


  Query Match             37.4%;  Score 594;  DB 22;  Length 245;
  Best Local Similarity   50.7%;  Pred. No. 1e-39;
  Matches  151;  Conservative   26;  Mismatches   63;  Indels   58;  Gaps   15;

Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASL-DCTVKSGESVVLSVTVGGAKLVIGTLSQ 59
           |||||: |       |||||  |:|:||||| :|  | || | | | ||   ||:|||
Db       1 mefwgievksgkpvtvtpeegilihvsqaslgecknkkgefvplhvkvgnqnlvlgtlst 60

Qy      60 DKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPT 119
           :  ||:  ||||||||||||:  | :|:|:|||:||||   : | ::|:  | |||
Db      61 enipqlfcdlvfdkefelshtwgkgsvyfvgyktpniepqgy-seeeeeeeeevpa---- 115

Qy     120 AVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMDVD 179
                 |||  ||     |||||||||||| :    ||||||| :           ||  |
Db     116 -----gnaakav----akpkakpaevkpavd----deedesdsd----------gm--d 149

Qy     180 EDDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTE 239
           |||||    |||  |||| |||||   :||| ||:  | ||| |||| |     |||||:

Db     150 eddsdg---eds--eeeeptpkkpas-skkranettpkapvsakkakvav----tpqktd 199

Qy     240 KKKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGGNNNKKPFNSGKQFGGSNNKGSN 297
           :|           |||||:   |||||||  | | | : || ||||     |:||  :
Db     200 ek----------kkggka---anqspksasqvscg-sckktfnsgnal-eshnkakh 241


RESULT   6
AAG31484
ID   AAG31484 standard; Protein; 257 AA.
XX
AC   AAG31484;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 37817.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
```

PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 37.4%; Score 594; DB 21; Length 257;
Best Local Similarity 50.7%; Pred. No. 1.1e-39;
Matches 151; Conservative 26; Mismatches 63; Indels 58; Gaps 15;

```
Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASL-DCTVKSGESVVLSVTVGGAKLVIGTLSQ 59
           |||||: |       |||||  |:|:||||| :|  | || | | | ||   ||:||||
Db       1 mefwgievksgkpvtvtpeegililhvsqaslgecknkkgefvplhvkvgnqnlvlgtlst 60

Qy      60 DKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPT 119
```

```
          :  ||:  ||||||||||||:  | :|:|:|||:||||   : | ::|:  | |||
Db     61 enipqlfcdlvfdkefelshtwgkgsvyfvgyktpniepqgy-seeeeeeeeevpa---- 115

Qy    120 AVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMDVD 179
              |||  ||     ||||||||||||| :    ||||||| :           ||  |
Db    116 -----gnaakav----akpkakpaevkpavd----deedesdsd-----------gm--d 149

Qy    180 EDDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTE 239
          |||||    |||  |||| |||||   :||| ||:  | ||| |||| |     |||||:
Db    150 eddsdg---eds--eeeeptpkkpas-skkranettpkapvsakkakvav----tpqktd 199

Qy    240 KKKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGGNNNKKPFNSGKQFGGSNNKGSN 297
          :|           |||||:   |||||||  | | } : || ||||     |:||  :
Db    200 ek-----------kkggka---anqspksasqvscg-sckktfnsgnal-eshnkakh 241


RESULT   7
AAG10894
ID   AAG10894 standard; Protein; 248 AA.
XX
AC   AAG10894;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 9401.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
PR   02-AUG-1999;   99US-0146386.
PR   02-AUG-1999;   99US-0146388.
PR   02-AUG-1999;   99US-0146389.
PR   03-AUG-1999;   99US-0147038.
PR   04-AUG-1999;   99US-0147204.
PR   04-AUG-1999;   99US-0147302.
```

PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```
  Query Match             37.1%;  Score 589.5;  DB 21;  Length 248;
  Best Local Similarity   49.7%;  Pred. No. 2.3e-39;
  Matches  148;  Conservative   27;  Mismatches   68;  Indels   55;  Gaps   14;

Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASL-DCTVKSGESVVLSVTVGGAKLVIGTLSQ 59
           |||||: |       |||||  |:|:||||| :|  | || | | | ||   ||:||||
Db       1 mefwglevksgkpvtvtpeegilihvsqaslgecknkkgefvplhvkvgnqnlvlgtlst 60

Qy      60 DKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPT 119
           :  ||:  |||||||||||:  | :|:|:|||:||||   ::  ::|:  |    ||
Db      61 enipqlfcdlvfdkefelshtwgkgsvyfvgyktpniepqgyseeeeeeeeeeevpa-- 118

Qy     120 AVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMDVD 179
               |||  ||     ||||||||||| :    ||||||| :           ||  |
Db     119 -----gnaakav----akpkakpaevkpavd---deedesdsd----------gm--d 152

Qy     180 EDDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTE 239
           |||||    |||  |||| |||||   :||| ||:  | ||| |||| |     |||||:
Db     153 eddsdg---eds--eeeeptpkkpas-skkranettpkapvsakkakvav----tpqktd 202

Qy     240 KKKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGGNNNKKPPNSGKQPGGSNNKGSN 297
           :|          |||||:   || ||||  | | | : || ||||     |:||  :
Db     203 ek----------kkggka---anxspksasqvscg-sckktfnsgnal-eshnkakh 244


RESULT   8
AAY28804
ID   AAY28804 standard; protein; 311 AA.
XX
AC   AAY28804;
XX
DT   13-JAN-2000  (first entry)
XX
DE   Maize histone deacetylase-8.
XX
KW   Maize histone deacetylase; HD; HD cDNA; family 2, ZmHD2; gene repression;
KW   acetyl modification; promoter; regulatory element; transgenic plant;
KW   disease resistance; toxin screening; pathogenicity;
KW   disease response promoter.
XX
OS   Zea mays.
XX
PN   WO9951731-A2.
XX
PD   14-OCT-1999.
XX
PF   02-APR-1999;   99WO-US07370.
XX
PR   03-APR-1998;   98US-0080563.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Baldwin DA,  Briggs SP,  Crane VC;
XX
DR   WPI; 1999-611038/52.
DR   N-PSDB; AAX90844.
XX
PT   New deacetylase genes, used for producing transgenic plants which have
PT   increased disease resistance  -
XX
PS   Claim 1; Page 79-80; 87pp; English.
XX
CC   The present sequence is maize histone deacetylase encoded by HD cDNA
CC   belonging to family 2, ZmHD2. This enzyme responsible for removing acetyl
CC   modifications, may be localised to promoters targeted for repression by
CC   other proteins that associate with HD and specifically bind regulatory
CC   elements in promoter DNA. The HD nucleotide sequence can be used for
CC   producing transgenic plants with increased disease resistance.
CC   Additionally, compositions find use in screening for toxins that affect
CC   pathogenicity and in determining which disease response promoters are
CC   regulated by histone deacetylase.
XX
SQ   Sequence   311 AA;
```

```
  Query Match             34.2%;  Score 544;  DB 20;  Length 311;
  Best Local Similarity   43.2%;  Pred. No. 1.3e-35;
  Matches  139;  Conservative   37;  Mismatches  104;  Indels   42;  Gaps   13;

Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQD 60
           |||||: | | :  |  |    ::|:|||:|  : || :| :: | |   || ||||| |
Db       1 mefwglevkpgstvkcepghgfilhvsqaalgeskks-dsalmyvkvddkklaigtlsid 59

Qy      61 KFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTA 120
           |:||| |||||:|||||||:    :| | |||     : |    | ||  | :  |
Db      60 kypqiqfdlvfnkefelshtsktsvffsgykveqpiegdemdldsedeeeelnip---v 116

Qy     121 VTANGNAGAAVVKADTKPKAKPAE--------------VKPAEEKPESDEEDESDDEDES 166
           :  ||       ||| | : |  |               | :::  :  :|||||| ||
Db     117 ikeng-------kadgkeeqknqekavaataskssiglekkskddsddsdedesddsded 169

Qy     167 EEDDDSE-KGMDVDEDDSDDDEEEDSEDEEEEE--TPKKPEPINKKRPNESVSKTPVSGK 223
           : ||  | :|:  || | |  :|:|: |::||| ||||||   ||| |:  |||:| |
Db     170 dsddsdegeglspdegdddssdeddtsdddeeetptpkkpe-agkkrgaenalktplsdk 228

Qy     224 KAKPAAAPASTPQKTEKKKGG-HTATPHPAKKGGKSPVN----ANQSPKSGGQSSGGNNN 278
           ||| |  ||   ||| ||| | ||||||| ||:| |      :|||||| |    :
Db     229 kakvatppa---qktggkkgathvatphpak--gktpanndksteksphsgg-svpcksc 282

Qy     279 KKPFNS--GKQFGGSNNKGSNK 298
           | |||    |     | | |:
Db     283 sktfnsemalqahskanmgpne 304
```

RESULT 9
AAY28803
ID AAY28803 standard; protein; 302 AA.
XX
AC AAY28803;
XX
DT 13-JAN-2000 (first entry)
XX
DE Maize histone deacetylase-7.
XX
KW Maize histone deacetylase; HD; HD cDNA; family 2, ZmHD2; gene repression;
KW acetyl modification; promoter; regulatory element; transgenic plant;
KW disease resistance; toxin screening; pathogenicity;
KW disease response promoter.
XX
OS Zea mays.
XX
PN WO9951731-A2.
XX
PD 14-OCT-1999.
XX
PF 02-APR-1999; 99WO-US07370.
XX
PR 03-APR-1998; 98US-0080563.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Baldwin DA, Briggs SP, Crane VC;
XX
DR WPI; 1999-611038/52.
DR N-PSDB; AAX90843.
XX
PT New deacetylase genes, used for producing transgenic plants which have
PT increased disease resistance -
XX
PS Claim 1; Page 76-77; 87pp; English.
XX
CC The present sequence is maize histone deacetylase encoded by HD cDNA
CC belonging to family 2, ZmHD2. This enzyme responsible for removing acetyl
CC modifications, may be localised to promoters targeted for repression by
CC other proteins that associate with HD and specifically bind regulatory
CC elements in promoter DNA. The HD nucleotide sequence can be used for
CC producing transgenic plants with increased disease resistance.
CC Additionally, compositions find use in screening for toxins that affect
CC pathogenicity and in determining which disease response promoters are
CC regulated by histone deacetylase.
XX
SQ Sequence 302 AA;

```
  Query Match             33.7%;  Score 535;  DB 20;  Length 302;
  Best Local Similarity   43.8%;  Pred. No. 6.6e-35;
  Matches  124;  Conservative   44;  Mismatches   93;  Indels   22;  Gaps    9;

Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQD 60
           |||||: | | :  |  |    ::|:|||:|  : || :: :: | :   || ||||| |
Db       1 mefwglevkpgstvkcepgygfvlhlsqaalgeskks-dnalmyvkiddqklaigtlsvd 59

Qy      61 KFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTA 120
           | | | |||:|||||||||:    :| | |||     ::|    | ||  | :  |
Db      60 knphiqfdlifdkefelshtsktsvfftgykveqpfeedemdldsededeelnvp---v 116

Qy     121 VTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEED--DDSEKGMDV 178
           |  || |     |:  |  | |::  |  :| : |:: : |: |:|:||  |||::|:
Db     117 vkengkadgkkqksqekavaapsksspdskkskddddsdedetddsdedetddsdeglss 176

Qy     179 DEDDSDDDEEEDSEDEEEEE--TPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQ 236
           :| | |  :|:|: |:|||:  ||||||   |:    || |||:| ||||  | |:|  |
Db     177 eegdddssdeddtsddeeedtptpkkpevgkkraaessvlktplsdkkak-vatpss--q 233

Qy     237 KTEKKKGG--HTATPHPAKKGGKSPVNANQS-------PKSGG 270
           ||  |||   | |||||||  ||: || ::|       |||||
Db     234 ktggkkgaavhvatphpak--gktivnndksvkspksapksgg 274
```

RESULT 10
AAY28802
ID AAY28802 standard; protein; 305 AA.
XX
AC AAY28802;
XX
DT 13-JAN-2000 (first entry)
XX
DE Maize histone deacetylase-6.
XX
KW Maize histone deacetylase; HD; HD cDNA; family 2, ZmHD2; gene repression;
KW acetyl modification; promoter; regulatory element; transgenic plant;
KW disease resistance; toxin screening; pathogenicity;
KW disease response promoter.
XX
OS Zea mays.
XX
PN WO9951731-A2.
XX
PD 14-OCT-1999.
XX
PF 02-APR-1999; 99WO-US07370.
XX
PR 03-APR-1998; 98US-0080563.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Baldwin DA, Briggs SP, Crane VC;
XX
DR WPI; 1999-611038/52.
DR N-PSDB; AAX90842.
XX
PT New deacetylase genes, used for producing transgenic plants which have
PT increased disease resistance -
XX
PS Claim 1; Page 72-73; 87pp; English.
XX
CC The present sequence is maize histone deacetylase encoded by HD cDNA
CC belonging to family 2, ZmHD2. This enzyme responsible for removing acetyl
CC modifications, may be localised to promoters targeted for repression by

```
CC   other proteins that associate with HD and specifically bind regulatory
CC   elements in promoter DNA. The HD nucleotide sequence can be used for
CC   producing transgenic plants with increased disease resistance.
CC   Additionally, compositions find use in screening for toxins that affect
CC   pathogenicity and in determining which disease response promoters are
CC   regulated by histone deacetylase.
XX
SQ   Sequence   305 AA;


  Query Match             33.2%;  Score 528;  DB 20;  Length 305;
  Best Local Similarity   43.6%;  Pred. No. 2.4e-34;
  Matches  123;  Conservative   44;  Mismatches   93;  Indels   22;  Gaps    9;

Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQD 60
           |||||: | | :  |  |    ::|:|||:|  : || :: :: | :   || ||||| |
Db       1 mefwglevkpgstvkcepgygfvihlsqaalgeskks-dnalmyvkiddqklaigtlsvd 59

Qy      61 KFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTA 120
           | | | |||:|||||||||:    :| | |||     ::|    | ||  | :  |
Db      60 knphiqfdlifdkefelshtskttsvfftgykveqpfeedemdldsededeelnvp---v 116

Qy     121 VTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDFSEED--DDSEKGMDV 178
           |  || |     |:  |  | |::  |  :| : |:: : |: |:|:||  |||::|:
Db     117 vkengkadekkqksqekavaapsksspdskkskddddsdedetddsdedetddsdeglsp 176

Qy     179 DEDDSDDDEEEDSEDEEEEE--TPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQ 236
           :| | |  :|:|: |:|||:  ||||||   |:     || |||:| ||||  | |:|  |
Db     177 eegdddssdeddtsddeeedtptpkkpevgkkraaessvlktplsdkkak-vatpss--q 233

Qy     237 KTEKKKGG--HTATPHPAKKGGKSPVNANQS-------PKSG 269
           ||  |||   | |||||||  ||: || ::|       ||||
Db     234 ktggkkgaavhvatphpak--gktivnndksvkspksapksg 273


RESULT  11
AAY28805
ID   AAY28805 standard; protein; 285 AA.
XX
AC   AAY28805;
XX
DT   13-JAN-2000  (first entry)
XX
DE   Maize histone deacetylase-9.
XX
KW   Maize histone deacetylase; HD; HD cDNA; family 2, ZmHD2; gene repression;
KW   acetyl modification; promoter; regulatory element; transgenic plant;
KW   disease resistance; toxin screening; pathogenicity;
KW   disease response promoter.
XX
OS   Zea mays.
XX
PN   WO9951731-A2.
XX
PD   14-OCT-1999.
XX
PF   02-APR-1999;   99WO-US07370.
XX
PR   03-APR-1998;   98US-0080563.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Baldwin DA,  Briggs SP,  Crane VC;
XX
DR   WPI; 1999-611038/52.
DR   N-PSDB; AAX90845.
XX
PT   New deacetylase genes, used for producing transgenic plants which have
PT   increased disease resistance  -
XX
PS   Claim 1; Page 82-84; 87pp; English.
XX
CC   The present sequence is maize histone deacetylase encoded by HD cDNA
CC   belonging to family 2, ZmHD2. This enzyme responsible for removing acetyl
CC   modifications, may be localised to promoters targeted for repression by
CC   other proteins that associate with HD and specifically bind regulatory
CC   elements in promoter DNA. The HD nucleotide sequence can be used for
CC   producing transgenic plants with increased disease resistance.
CC   Additionally, compositions find use in screening for toxins that affect
CC   pathogenicity and in determining which disease response promoters are
CC   regulated by histone deacetylase.
XX
SQ   Sequence   285 AA;


  Query Match             27.4%;  Score 436;  DB 20;  Length 285;
  Best Local Similarity   36.0%;  Pred. No. 5e-27;
  Matches  114;  Conservative   49;  Mismatches  110;  Indels   44;  Gaps   11;

Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQD 60
           |||||  | |          :  ::|:|||:|    |: |: ::|| :   |||:|||| :
Db       1 mefwgeevkpgatvsckvgdglvihlsqaalgepkkasenaivsvkiddkklvlgtlsve 60

Qy      61 KFPQISFDLVFDKEFELSHSGTKANVHFIGYKS--PNIEQDDFTSSDDEDV-PEAVPAPA 117
           | |||| ||||||:|||||:   |:| | ||||  |  | |     | ||:| |: :|
Db      61 khpqiscdlvfdkdfelshnsktasvffcgykspvplfesdsgedssdeevepdlip--- 117

Qy     118 PTAVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMD 177
                 |     :  |   |   : |:    :|    |::| :| : |  |:|:|    |
Db     118 ----mqnneikistakvpvkvgiqnad---edetssgddddftdsdsemseedess---d 167

Qy     178 VDEDDSDDDEEEDSEDEEEEETPKKPE-PINKKRPNESVSKTPVSGKKAKPAAAPASTPQ 236
             ||  || |  :||  ||:  |||| |  : |||  |  ::|| ||||||        :
Db     168 edevssdtdtsddsgseeqtptpkktevvvgkkraie--aetp-sgkkak--------se 216

Qy     237 KTEKKKGGHTATPHPAKKGGKSPVNAN--------QSPKSGGQSSGGNNNKKPPNSGKQP 288
           :: :| |   :| ||||:  |:| : :         :|||||  :          :  | |
Db     217 qsaqktgdkvstshpakqssktpadkstktptadkkspksgshack--------scsksf 268

Qy     289 GGSNNKGSNKGKGKGRA 305
           | ::   |::   |  |
Db     269 gsasaleshqkakkhea 285


RESULT  12
AAG54981
ID   AAG54981 standard; Protein; 149 AA.
XX
AC   AAG54981;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 70363.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
```

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 24.8%; Score 393.5; DB 21; Length 149;
Best Local Similarity 59.4%; Pred. No. 5.8e-24;
Matches 82; Conservative 5; Mismatches 14; Indels 37; Gaps 2;

```
Qy    205 PINKKRPNESVSKTPVSGKKAKPA---------------------------------- 228
          |:|: | :  :   |:   || |
Db     12 plngtrsspflkplplkpskalvatggraqrlqvkalkmdkaltgisaaaltasmvipei 71

Qy    229 AAPASTPQKT-EKKKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGGNNNKKPFNSGKQ 287
          |  ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     72 aeaastpqkteckkkgghtatphpakkggkspvnanqspksggqssggnnnkkpfnsgkq 131

Qy    288 FGGSNNKGSNKGKGKGRA 305
          ||||||||||||||||||
Db    132 fggsnnkgsnkgkgkgra 149
```

RESULT 13
AAG25823
ID AAG25823 standard; Protein; 208 AA.
XX
AC AAG25823;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 30043.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.7%; Score 345.5; DB 21; Length 208;
Best Local Similarity 39.6%; Pred. No. 5.8e-20;
Matches 93; Conservative 27; Mismatches 66; Indels 49; Gaps 9;

```
Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGES------VVLSVTVGGAKLV 53
           |||||| |       : |  | |||||| :|      |||       : | ||||  ||:
Db       1 mefwgvevkngkplhldpgldrlvhisqval------geskpnvtepiqlyvtvgsdkll 54

Qy      54 IGTLSQDKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAV 113
           ||||| :||||:| ::| :: | |||:    :| | |||       | :  : ||:
Db      55 igtlshekfpqlsteivlernfalshtwkngsvffsgykv------dasdpepedl---- 104

Qy     114 PAPAPTAVTANGNAGAAVVKADTKPKAK------PAEVKPAEEKPESD--EEDESDDEDE 165
                   :    ||  ||  |  ||      | |   |::  ::|  ||| |||:|
Db     105 ---------iddqleaagfkaapksaakqvnfqlpnedvkakqdddadgseedssdddds 155

Qy     166 SEEDDDSEKGM----DVDEDDSDDDEEEDSEDEEEEETPKKPEPINKK--RPNES 214
             |: |: :    | :|||| ||||:||    |||||||||   |:   || |
Db     156 ensgdeeeekvtaesdseeddssddeedds---seeetpkkpeepkkrsaepnss 207
```

RESULT 14
AAW68010
ID AAW68010 standard; Protein; 412 AA.
XX
AC AAW68010;
XX
DT 27-APR-1999 (first entry)
XX
DE S. frugiperda immunophilin FKBP46.
XX
KW Immunophilin; moth; insect cell; nuclear; immunosuppression; drug;
KW transplant; tissue graft.
XX
OS Spodoptera frugiperda.
XX

FH	Key	Location/Qualifiers
FT	Domain	111..114
FT		/note= "EEAP motif"
FT	Region	124
FT		/note= "putative nuclear localisation signal"
FT	Region	135..138
FT		/note= "putative nuclear localisation signal"
FT	Domain	148..149
FT		/note= "putative nuclear localisation signal"
FT		/note= "AP motif"
FT	Domain	215..218

```
FT                   /note= "EEAP motif"
FT   Region          219..223
FT                   /note= "putative nuclear localisation signal"
FT   Region          250..256
FT                   /note= "putative nuclear localisation signal"
FT   Region          272..275
FT                   /note= "putative nuclear localisation signal"
FT   Region          284..286
FT                   /note= "putative nuclear localisation signal"
FT   Domain          292..295
FT                   /note= "EEAP motif"
FT   Region          298..302
FT                   /note= "putative nuclear localisation signal"
XX
PN   US5861498-A.
XX
PD   19-JAN-1999.
XX
PF   31-OCT-1996;   96US-0741134.
XX
PR   01-NOV-1995;   95US-0007163.
PR   31-OCT-1996;   96US-0741134.
XX
PA   (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI   Alnemri ES,  Fernandes-Alnemri T,  Litwack G;
XX
DR   WPI; 1999-130433/11.
DR   N-PSDB; AAX04441.
XX
PT   Isolated immunophilin FKBP46 nucleic acids - useful for developing
PT   products for the study and identification of immunosuppressive
PT   agents for treating e.g. transplant and tissue graft patients
XX
PS   Claim 1; Fig 1; 21pp; English.
XX
CC   This sequence represents the immunophilin FKBP46 from the moth Spodoptera
CC   frugiperda.  FKBP46 is a novel insect cell nuclear immunophilin, which
CC   can be used to study and identify additional immunosuppressive drugs that
CC   bind to it.  Such immunosuppressive agents can be used in treating
CC   transplant and tissue graft patients.
XX
SQ   Sequence   412 AA;


  Query Match             19.6%;  Score 311;  DB 20;  Length 412;
  Best Local Similarity   31.0%;  Pred. No. 7.3e-17;
  Matches  100;  Conservative   39;  Mismatches  116;  Indels   68;  Gaps   14;

Qy       3 FWGVAVTP-KNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQDK 61
           |||: : | |  |:|    :   |||||::| :    :   : | | |   :: || : |
Db       2 fwglimepnkrytqvv---ekpfhisqaamdistgdndpcqvmvvvdgknflvctlqkgk 58

Qy      62 FPQISFDLVFDKEFELSH-SGTKANVHFIGYKSPNIEQD--DFTSSDDEDVPEAVPAPAP 118
             |:  || |     :|  :  | |||  ||  |  |:|  |   :::|:  |  |   |
Db      59 iiqvpldlyfksgdsvsfltngkcnvhltqyldpefeedledeeeaeeeeeeeeapplvp 118

Qy     119 T---AVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEED----------ESDDED 164
               | |   |  ||    ||       ||   : |:||           ::|: |
Db     119 aknkrklenandatankkakpdkkagknsapaaesdsddddedqlqkfldqedidtdend 178

Qy     165 ESEEDDDSEKGMDVDEDDSDDDEE----EDSEDEEEEETPKKPEPINKKRP--------- 211
           || : : | :| | ||:| |:|||    :| :||||| |||     ||:|
Db     179 esfkmntsaegddsdeedddedeedeedddeddeeeeeapkk----kkkqpaaeqdstld 234

Qy     212 --NESV--SKTPVSGK----------KAKPAAAPASTP------QKTEKKKGGHTATPH 250
             |||  ||   | |           | :|       |      || ||||
Db     235 tskesvdmsklsksqkrrlkkklqqqakqqpqvngvdkpkkeepqqkaekkk-------- 286

Qy     251 PAKKGGKSPVNANQSPK-SGGQS 272
           |  |  ::||   :  : :|| |
Db     287 peakkeeapvekkekkqiaggvs 309
```

```
RESULT  15
AAG09171
ID   AAG09171 standard; Protein; 203 AA.
XX
AC   AAG09171;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 6998.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
```

PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 16.5%; Score 262.5; DB 21; Length 203;
Best Local Similarity 30.9%; Pred. No. 2.4e-13;
Matches 80; Conservative 33; Mismatches 75; Indels 71; Gaps 11;

```
Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCT--VKSGESVVLSVTVGGAK--LVIGT 56
           |||||: : |    ||  ::  :|| || :|     ||  |:  : | :|  :   :||
Db       1 mefwqieikpgkpfkviqkdgfmvhasqvtlgdvekvkkdetfavyvkigddengfmign 60

Qy      57 LSQDKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNI--EQDDFTSSDDEDVPEAVP 114
           ||| |||| | ||    |||:||: | ::|: |||:: :   | |:   || |
Db      61 lsq-kfpqfsidlylghefeishnst-ssvyligyrtfdafdeldeeidsdse------- 111

Qy     115 APAPTAVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEK 174
                           |   : : | :   |  || |:| :||           |
Db     112 -------------------ldeymeqqiaalpqneinpeeddesdsd----------em 141
```

```
Qy    175 GMDVDEDDSDDDEEEDSEDEEEEETPKKPEPINKKRPNES--VSKTPVSGKKAKPAAAPA 232
          |:| |:| ||       |:: | | | |  | :|| || :  ::|    |||
Db    142 gldedddssd-------eedveaeaplkvappskkmpngafeiak---ggkk-------- 183

Qy    233 STPQKTEKKKGGHTATPHP 251
              |  ||    | |
Db    184 ------nkssqgkkrcpfp 196
```

Search completed: May 2, 2002, 22:03:53
Job time: 6541 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: May 2, 2002, 22:08:41 ; Search time 60.74 Seconds
(without alignments)
122.651 Million cell updates/sec

Title: US-09-645-337-8
Perfect score: 1589
Sequence: 1 MEFWGVAVTPKNATKVTPEE.....KQFGGSNNKGSNKGKGKGRA 305

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	544	34.2	311	4	US-09-282-305-16	Sequence 16, Appl
2	535	33.7	302	4	US-09-282-305-14	Sequence 14, Appl
3	528	33.2	305	4	US-09-282-305-12	Sequence 12, Appl
4	436	27.4	285	4	US-09-282-305-18	Sequence 18, Appl
5	311	19.6	412	2	US-08-741-134-2	Sequence 2, Appli
6	206	13.0	714	2	US-08-990-114-3	Sequence 3, Appli
7	206	13.0	714	4	US-09-241-333-3	Sequence 3, Appli
8	188	11.8	723	1	US-07-814-964-11	Sequence 11, Appl
9	188	11.8	723	1	US-08-258-442-11	Sequence 11, Appl
10	188	11.8	723	1	US-08-328-809-6	Sequence 6, Appli
11	188	11.8	723	5	PCT-US92-11107-11	Sequence 11, Appl
12	180	11.3	1162	2	US-08-728-323A-2	Sequence 2, Appli
13	170	10.7	764	1	US-08-375-300-4	Sequence 4, Appli
14	170	10.7	764	3	US-09-177-431-4	Sequence 4, Appli
15	170	10.7	764	5	PCT-US95-16930-4	Sequence 4, Appli
16	170	10.7	1089	1	US-08-375-300-2	Sequence 2, Appli
17	170	10.7	1089	3	US-09-177-431-2	Sequence 2, Appli
18	170	10.7	1089	5	PCT-US95-16930-2	Sequence 2, Appli
19	169	10.6	292	4	US-09-538-871-2	Sequence 2, Appli
20	168.5	10.6	739	4	US-09-022-983-2	Sequence 2, Appli
21	168	10.6	411	2	US-08-741-134-6	Sequence 6, Appli
22	167.5	10.5	740	4	US-09-022-983-5	Sequence 5, Appli
23	166.5	10.5	257	4	US-09-538-871-4	Sequence 4, Appli
24	164	10.3	542	1	US-08-701-380-2	Sequence 2, Appli
25	164	10.3	542	3	US-09-032-365A-13	Sequence 13, Appl
26	161.5	10.2	905	2	US-08-574-959A-9	Sequence 9, Appli
27	161.5	10.2	905	4	US-09-357-014-9	Sequence 9, Appli
28	161.5	10.2	1135	2	US-08-574-959A-7	Sequence 7, Appli
29	161.5	10.2	1135	4	US-09-357-014-7	Sequence 7, Appli
30	156	9.8	1085	1	US-08-431-080-28	Sequence 28, Appl
31	156	9.8	1085	2	US-08-938-534-28	Sequence 28, Appl
32	155.5	9.8	1187	1	US-08-320-559-28	Sequence 28, Appl
33	155.5	9.8	1187	3	US-08-545-860D-28	Sequence 28, Appl
34	155.5	9.8	1187	5	PCT-US94-04496-28	Sequence 28, Appl
35	155.5	9.8	1210	1	US-08-320-559-26	Sequence 26, Appl
36	155.5	9.8	1210	3	US-08-545-860D-26	Sequence 26, Appl
37	155.5	9.8	1210	5	PCT-US94-04496-26	Sequence 26, Appl
38	149	9.4	231	4	US-09-461-697-194	Sequence 194, App
39	149	9.4	232	4	US-09-461-697-192	Sequence 192, App
40	149	9.4	238	4	US-09-461-697-190	Sequence 190, App
41	149	9.4	257	4	US-09-461-697-188	Sequence 188, App
42	149	9.4	272	4	US-09-461-697-186	Sequence 186, App
43	149	9.4	546	3	US-08-935-855-20	Sequence 20, Appl
44	148.5	9.3	251	2	US-08-766-738-3	Sequence 3, Appli
45	146	9.2	160	2	US-08-726-306A-68	Sequence 68, Appl

ALIGNMENTS

RESULT 1
US-09-282-305-16
; Sequence 16, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
; APPLICANT: Baldwin, Donald A.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Crane, Virginia C.
; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
; FILE REFERENCE: 5718-44,
; CURRENT APPLICATION NUMBER: US/09/282,305
; CURRENT FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 60/080,563
; PRIOR FILING DATE: 1998-04-03
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Zea mays
US-09-282-305-16

Query Match 34.2%; Score 544; DB 4; Length 311;
Best Local Similarity 43.2%; Pred. No. 6.8e-39;
Matches 139; Conservative 37; Mismatches 104; Indels 42; Gaps 13;

```
Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQD 60
           |||||: | | :  |  |    ::|:|||:|  : || :| :: | |   || ||||| |
Db       1 MEFWGLEVKPGSTVKCEPGHGFILHVSQAALGESKKS-DSALMYVKVDDKKLAIGTLSID 59

Qy      61 KFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTA 120
           |:||| |||||:|||||||:    :| | |||     : |   | ||  | :  |
Db      60 KYPQIQFDLVFNKEFELSHTSKTTSVFFSGYKVEQPIEGDEMDLDSEDEEEELNIP---V 116

Qy     121 VTANGNAGAAVVKADTKPKAKPAE-------------VKPAEEKPESDEEDESDDEDES 166
           :  ||       ||| | : |  |                | :::  :  :|||||| ||
Db     117 IKENG-------KADGKEEQKNQEKAVAATASKSSLGLEKKSKDDSDDSDEDESDDSDED 169

Qy     167 EEDDDSE-KGMDVDEDDSDDDEEEDSEDEEEEE--TPKKPEPINKKRPNESVSKTPVSGK 223
           : ||  | :|:  || | |  :|:|: |::|||  ||||||   |||  |:  |||:| |
Db     170 DSDDSDEGEGLSPDEGDDDSSDEDDTSDDDEEETPTPKKPE-AGKKRGAENALKTPLSDK 228

Qy     224 KAKPAAAPASTPQKTEKKKGG-HTATPHPAKKGGKSPVN----ANQSPKSGGQSSGGNNN 278
           ||| |  ||   |||  |||  | |||||||  ||:| |      :||||| |    :
Db     229 KAKVATPPA---QKTGGKKGATHVATPHPAK--GKTPANNDKSTEKSPKSGG-SVPCKSC 282

Qy     279 KKPFNS--GKQFGGSNNKGSNK 298
           | |||    |     | | |:
Db     283 SKTFNSEMALQAHSKANMGPNE 304
```

```
RESULT   2
US-09-282-305-14
; Sequence 14, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
;  APPLICANT:  Baldwin, Donald A.
;  APPLICANT:  Briggs, Steven P.
;  APPLICANT:  Crane, Virginia C.
;  TITLE OF INVENTION:  Maize Histone Deacetylases And Their Uses
;  FILE REFERENCE:  5718-44,
;  CURRENT APPLICATION NUMBER: US/09/282,305
;  CURRENT FILING DATE:  1999-03-31
;  PRIOR APPLICATION NUMBER:  60/080,563
;  PRIOR FILING DATE:  1998-04-03
;  NUMBER OF SEQ ID NOS:  18
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 302
;   TYPE: PRT
;   ORGANISM: Zea mays
US-09-282-305-14
```

```
  Query Match             33.7%;  Score 535;  DB 4;  Length 302;
  Best Local Similarity   43.8%;  Pred. No. 3.8e-38;
  Matches  124;  Conservative   44;  Mismatches   93;  Indels   22;  Gaps    9;

Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQD 60
           |||||: | | :  |  |    ::|:|||:|  : || :: :: | :   || ||||| |
Db       1 MEFWGLEVKPGSTVKCEPGYGFVLHLSQAALGESKKS-DNALMYVKIDDQKLAIGTLSVD 59

Qy      61 KFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTA 120
           | | | |||:|||||||||:    :| | |||     ::|    | ||  | :  |
Db      60 KNPHIQFDLIFDKEFELSHTSKTTSVFFTGYKVEQPFEEDEMDLDSEDEDEELNVP---V 116

Qy     121 VTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEED--DDSEKGMDV 178
           |  || |      |:  |  | |::  |  :| : |:: : |: |:|:||  |||::|:
Db     117 VKENGKADKKQKSQEKAVAAPSKSSPDSKKSKDDDDSDEDETDDSDEDETDDSDEGLSS 176

Qy     179 DEDDSDDDEEEDSEDEEEEE--TPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQ 236
           :| | |  :|:|: |:|||:  ||||||   |:    || |||:| |||| | |:|  |
Db     177 EEGDDDSSDEDDTSDDEEEDTPTPKKPEVGKKRAAESSVLKTPLSDKKAK-VATPSS--Q 233

Qy     237 KTEKKKGG--HTATPHPAKKGGKSPVNANQS-------PKSGG 270
           ||  |||   | |||||||  ||: || ::|       |||||
Db     234 KTGGKKGAAVHVATPHPAK--GKTIVNNDKSVKSPKSAPKSGG 274
```

```
RESULT   3
US-09-282-305-12
; Sequence 12, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
;  APPLICANT:  Baldwin, Donald A.
;  APPLICANT:  Briggs, Steven P.
;  APPLICANT:  Crane, Virginia C.
;  TITLE OF INVENTION:  Maize Histone Deacetylases And Their Uses
;  FILE REFERENCE:  5718-44,
;  CURRENT APPLICATION NUMBER: US/09/282,305
;  CURRENT FILING DATE:  1999-03-31
;  PRIOR APPLICATION NUMBER:  60/080,563
;  PRIOR FILING DATE:  1998-04-03
;  NUMBER OF SEQ ID NOS:  18
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 12
;   LENGTH: 305
;   TYPE: PRT
;   ORGANISM: Zea mays
US-09-282-305-12
```

```
  Query Match             33.2%;  Score 528;  DB 4;  Length 305;
  Best Local Similarity   43.6%;  Pred. No. 1.5e-37;
  Matches  123;  Conservative   44;  Mismatches   93;  Indels   22;  Gaps    9;

Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQD 60
           |||||: | | :  |  |    ::|:|||:|  : || :: :: | :   || ||||| |
Db       1 MEFWGLEVKPGSTVKCEPGYGFVLHLSQAALGESKKS-DNALMYVKIDDQKLAIGTLSVD 59

Qy      61 KFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTA 120
           | | | |||:|||||||||:    :| | |||     ::|    | ||  | :  |
Db      60 KNPHIQFDLIFDKEFELSHTSKTTSVFFTGYKVEQPFEEDEMDLDSEDEDEELNVP---V 116

Qy     121 VTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEED--DDSEKGMDV 178
           |  || |      |:  |  | |::  |  :| : |:: : |: |:|:||  |||::|:
Db     117 VKENGKADEKKQKSQEKAVAAPSKSSPDSKKSKDDDDSDEDETDDSDEDETDDSDEGLSP 176

Qy     179 DEDDSDDDEEEDSEDEEEEE--TPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQ 236
           :| | |  :|:|: |:|||:  ||||||   |:    || |||:| |||| | |:|  |
Db     177 EEGDDDSSDEDDTSDDEEEDTPTPKKPEVGKKRAAESSVLKTPLSDKKAK-VATPSS--Q 233

Qy     237 KTEKKKGG--HTATPHPAKKGGKSPVNANQS-------PKSG 269
           ||  |||   | |||||||  ||: || ::|       ||||
Db     234 KTGGKKGAAVHVATPHPAK--GKTIVNNDKSVKSPKSAPKSG 273
```

```
RESULT   4
US-09-282-305-18
; Sequence 18, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
;  APPLICANT:  Baldwin, Donald A.
;  APPLICANT:  Briggs, Steven P.
;  APPLICANT:  Crane, Virginia C.
;  TITLE OF INVENTION:  Maize Histone Deacetylases And Their Uses
;  FILE REFERENCE:  5718-44,
;  CURRENT APPLICATION NUMBER: US/09/282,305
;  CURRENT FILING DATE:  1999-03-31
;  PRIOR APPLICATION NUMBER:  60/080,563
;  PRIOR FILING DATE:  1998-04-03
;  NUMBER OF SEQ ID NOS:  18
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 18
;   LENGTH: 285
;   TYPE: PRT
;   ORGANISM: Zea mays
US-09-282-305-18
```

```
  Query Match             27.4%;  Score 436;  DB 4;  Length 285;
  Best Local Similarity   36.0%;  Pred. No. 9.6e-30;
  Matches  114;  Conservative   49;  Mismatches  110;  Indels   44;  Gaps   11;

Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQD 60
           |||||  | |          :  ::|:|||:|    |: |: ::|| :   |||:|||| :
Db       1 MEFWGEEVKPGATVSCKVGDGLVIHLSQAALGEPKKASENAIVSVKIDDKKLVLGTLSVE 60

Qy      61 KFPQISFDLVFDKEFELSHSGTKANVHFIGYKS--PNIEQDDFTSSDDEDV-PEAVPAPA 117
           | |||| ||||||:|||||:    |:| | |||| |  | |    | ||:| |: :|
Db      61 KHPQISCDLVFDKDFELSHNSKTASVFFCGYKSPVPLFESDSGEDSSDEEVEPDLIP--- 117

Qy     118 PTAVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMD 177
           |      :  |   |  : |:    :|    |::| :| : |  |:|:|    |
Db     118 ----MQNNEIKISTAKVPVKVGIQNAD---EDETSSGDDDDFTDSDSEMSEEDESS---D 167

Qy     178 VDEDDSDDDEEEDSEDEEEEETPKKPE-PINKKRPNESVSKTPVSGKKAKPAAAPASTPQ 236
           ||  || |  :||  ||:  |||| |  : |||  |  ::|| ||||||         :
Db     168 EDEVSSDTDTSDDSGSEEQTPTPKKTEVVVGKKRAIE--AETP-SGKKAK--------SE 216

Qy     237 KTEKKKGGHTATPHPAKKGGKSPVNAN--------QSPKSGGQSSGGNNNKKPFNSGKQF 288
           :: :| |   :| ||||:  |:| : :         :|||||  :           : | |
Db     217 QSAQKTGDKVSTSHPAKQSSKTPADKSTKTPTADKKSPKSGSHACK--------SCSKSF 268

Qy     289 GGSNNKGSNKGKGKGRA 305
```

```
Db    269 GSASALESHQKAKKHEA 285


RESULT   5
US-08-741-134-2
; Sequence 2, Application US/08741134
; Patent No. 5861498
;  GENERAL INFORMATION:
;    APPLICANT:  Litwack, Gerald
;    APPLICANT:  Alnemri, Emad S.
;    APPLICANT:  Fernandes-Alnemri, Teresa
;    TITLE OF INVENTION:  IMMUNOPHILIN FKBP46 AND COMPOSITIONS FOR MAKING
;    TITLE OF INVENTION:  AND
;    TITLE OF INVENTION:  METHODS OF USING THE SAME
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz & No. 5861498ris
;      STREET:  One Liberty Place - 46th floor
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 3.11
;      SOFTWARE:  WordPerfect for Windows 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/741,134
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/007,163
;      FILING DATE:  01-NOV-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DeLuca, Mark
;      REGISTRATION NUMBER:  33,229
;      REFERENCE/DOCKET NUMBER:  TJU-2090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH: 412 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-741-134-2


  Query Match             19.6%;  Score 311;  DB 2;  Length 412;
  Best Local Similarity   31.0%;  Pred. No. 6.8e-19;
  Matches  100;  Conservative   39;  Mismatches  116;  Indels   68;  Gaps   14;

Qy       3 FWGVAVTP-KNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQDK 61
Db       2 FWGLIMEPNKRYTQVV---EKPFHISQAAMDISTGDNDPCQVMVVVDGKNFLVCTLQKGK 58

Qy      62 FPQISFDLVFDKEFELSH-SGTKANVHFIGYKSPNIEQD--DFTSSDDEDVPEAVPAPAP 118
Db      59 IIQVPLDLYFKSGDSVSFLTNGKCNVHLTGYLDPEFEEDLEDEEEAEEEEEEEEAPPLVP 118

Qy     119 T---AVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEED----------ESDDED 164
Db     119 AKNKRKLENANDATANKKAKPDKKAGKNSAPAAESDSDDDDEDQLQKFLDGEDIDTDEND 178

Qy     165 ESEEDDDSEKGMDVDEDDSDDDEE----EDSEDEEEEETPKKPEPINKKRP--------- 211
Db     179 ESFKMNTSAEGDDSDEEDDDEDEEDEEDDDDEDDEEEEEAPKK----KKKQPAAEQDSTLD 234

Qy     212 --NESV--SKTPVSGK----------KAKPAAAPASTP------QKTEKKKGGHTATPH 250
Db     235 TSKESVDMSKLSKSQKRRLKKKLQQQAKQQPQVNGVDKPKKEEPQQKAEKKK-------- 286

Qy     251 PAKKGGKSPVNANQSPK-SGGQS 272
Db     287 PEAKKEEAPVEKKEKKQIAGGVS 309


RESULT   6
US-08-990-114-3
; Sequence 3, Application US/08990114
; Patent No. 5932475
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  HUMAN NUCLEOLIN-LIKE PROTEIN
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/990,114
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0451 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  714 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  128842
US-08-990-114-3


  Query Match             13.0%;  Score 206;  DB 2;  Length 714;
  Best Local Similarity   23.2%;  Pred. No. 1.2e-09;
  Matches   66;  Conservative   31;  Mismatches   81;  Indels  106;  Gaps    8;

Qy      97 EQDDFTSSDDEDVPE----------AVPAPA---------- 117
Db      35 EEEDDSSGEEVVIPQKKGKKATATPAKKVVVSQTKKVAVPTPAKKAAVTPGKKAAATPAK 94

Qy     118 -----PTAVTANGNAGAAVVKADTKPKAKPAEVKPAE-------EKPESDEEDESDDEDE 165
Db      95 KAVTPAKAVATPGKKGATQAKALVATPGKKGAVTPAKGAKNGKNAKKEDSDEDEDDDDDE 154

Qy     166 SEEDDDSEKGMDVDEDD----------SDDDEEEDSEDEEEEETPKKPEP 205
Db     155 DDSDEDEE---DEEEDEFEPPVVKGKQGKVAAAAPASEDEDEEEDEEEEEEDE-----EE 206
```

```
Qy     206 INKKRPNESVSKTPVSGKKAKPAAAPAST-------------------------PQKTE 239
Db     207 EDDSEEEEAMEITPAKGKKAPAKVVPVKAKNVAEEDDDEEEDEDEEEDEEEEDEEEEE 266

Qy     240 KKKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGGNNNKKPFN 283
Db     267 EEEEEEPVKPAPGKR-KKEMTKQKEVPEAKKQKVEGSESTTPFN 309


RESULT   7
US-09-241-333-3
; Sequence 3, Application US/09241333
; Patent No. 6313266
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  HUMAN NUCLEOLIN-LIKE PROTEIN
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/241,333
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/990,114
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0451 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  714 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  128842
US-09-241-333-3


  Query Match             13.0%;  Score 206;  DB 4;  Length 714;
  Best Local Similarity   23.2%;  Pred. No. 1.2e-09;
  Matches   66;  Conservative   31;  Mismatches   81;  Indels  106;  Gaps    8;

Qy      97 EQDDFTSSDDEDVPE---------------------AVPAPA----------------- 117
Db      35 EEEDDSSGEEVVIPQKKGKKATATPAKKVVVSQTKKVAVPTPAKKAAVTPGKKAAATPAK 94

Qy     118 -----PTAVTANGNAGAAVVKADTKPKAKPAEVKPAE-------EKPESDEEDESDDEDE 165
Db      95 KAVTPAKAVATPGKKGATQAKALVATPGKKGAVTPAKGAKNGKNAKKEDSDEDEDDDDDE 154

Qy     166 SEEDDDSEKGMDVDEDD-------------------SDDDEEEDSEDEEEEETPKKPEP 205
```

```
Db     155 DDSDEDEE---DEEEDEFEPPVVKGKQGKVAAAAPASEDEDEEEDEEEEEEDE-----EE 206

Qy     206 INKKRPNESVSKTPVSGKKAKPAAAPAST-------------------------PQKTE 239
Db     207 EDDSEEEEAMEITPAKGKKAPAKVVPVKAKNVAEEDDDEEEDEDEEEDEEEEDEEEEE 266

Qy     240 KKKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGGNNNKKPFN 283
Db     267 EEEEEEPVKPAPGKR-KKEMTKQKEVPEAKKQKVEGSESTTPFN 309


RESULT   8
US-07-814-964-11
; Sequence 11, Application US/07814964
; Patent No. 5359047
;  GENERAL INFORMATION:
;    APPLICANT:  Donahue, Brian A.
;    APPLICANT:  Toney, Jeffrey H.
;    APPLICANT:  Bruhn, Suzanne L.
;    APPLICANT:  Pil, Pieter M.
;    APPLICANT:  Brown, Steven
;    APPLICANT:  Kellett, Patti
;    APPLICANT:  Essigmann, John M.
;    APPLICANT:  Lippard, Stephen J.
;    TITLE OF INVENTION:  DNA Structure Specific Recognition
;    TITLE OF INVENTION:  Protein and Uses Therefor
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  2 Militia Drive
;      CITY:  Lexington
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/814,964
;      FILING DATE:  19911226
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/539,906
;      FILING DATE:  18-JUN-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  MIT-4787AAA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  11
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  723 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Drosophila melanogaster
;    IMMEDIATE SOURCE:
;      CLONE:  Drosophila SSRP (predicted)
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  458..507
;      OTHER INFORMATION:  /label= Acidic
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  518..547
;      OTHER INFORMATION:  /label= Basic I
```

```
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  547..620
;      OTHER INFORMATION:  /label= HMG-box
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  632..649
;      OTHER INFORMATION:  /label= Basic II
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  657..723
;      OTHER INFORMATION:  /label= Mixed Charge
US-07-814-964-11

  Query Match             11.8%;  Score 188;  DB 1;  Length 723;
  Best Local Similarity   23.5%;  Pred. No. 4.2e-08;
  Matches   76;  Conservative   53;  Mismatches  110;  Indels   84;  Gaps   14;

Qy      19 EEDSLVHISQASLDCTVKSGESVVLSVTV-GGAKLVIGTLSQDKFPQISFDLVFDKEFEL 77
           || | |: :::       |  |    ||:  |   :  :: :::: :: || :  |:  :
Db     372 EEISSVNFARSG-----GSTRSFDFEVTLKNGTVHIFSSIEKEEYAKL-FDYITQKKLHV 425

Qy      78 SHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTAVTANGNAGAAVVKADTK 137
           |: |   :   |||       ||  ||:|: |:|                 | :||: :
Db     426 SNMGKDKS----GYKDV-----DFGDSDNENEPDAY--------------LARLKAEAR 461

Qy     138 PKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMDVD---EDDSDDDE------- 187
           |        |:  : | ::|| |||    :::|:   : |   | |||||
Db     462 EK--------EEDDDDGDSDEESTDEDFKPNENESDVAEEYDSNVESDSDDDSDASGGGG 513

Qy     188 EEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKK--AKPAAAPAS-------TPQKT 238
           : |   :::|:  :| |   ||   :  :| |   ||   ||  |  :       | :
Db     514 DSDGAKKKKEKKSEKKEKKEKKHKEKERTKKPSKKKKDSGKPKRATTAFMLWLNDTRESI 573

Qy     239 EKKKGGHTATPHPAKKGGK----------------------SPVNANQSPKSGGQS---S 273
           :::  |   |  |||||:                        |  |::|| |
Db     574 KRENPGIKVT-EIAKKGGEMWKELKDKSKWEDAAAKDKQRYHDEMRNYKPEAGGDSDNEK 632

Qy     274 GGNNNKKPFNSGKQFGGSNNKGS 296
           || ::||          :|  ||
Db     633 GGKSSKKRKTEPSPSKKANTSGS 655


RESULT   9
US-08-258-442-11
; Sequence 11, Application US/08258442
; Patent No. 5670621
;  GENERAL INFORMATION:
;    APPLICANT:  Donahue, Brian A.
;    APPLICANT:  Toney, Jeffrey H.
;    APPLICANT:  Bruhn, Suzanne L.
;    APPLICANT:  Pil, Pieter M.
;    APPLICANT:  Brown, Steven
;    APPLICANT:  Kellett, Patti
;    APPLICANT:  Essigmann, John M.
;    APPLICANT:  Lippard, Stephen J.
;    TITLE OF INVENTION:  DNA Structure Specific Recognition
;    TITLE OF INVENTION:  Protein and Uses Therefor
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  2 Militia Drive
;      CITY:  Lexington
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/258,442
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/539,906
;      FILING DATE:  18-JUN-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  MIT-4787AAA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  11:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  723 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Drosophila melanogaster
;    IMMEDIATE SOURCE:
;      CLONE:  Drosophila SSRP (predicted)
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  458..507
;      OTHER INFORMATION:  /label= Acidic
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  518..547
;      OTHER INFORMATION:  /label= Basic I
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  547..620
;      OTHER INFORMATION:  /label= HMG-box
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  632..649
;      OTHER INFORMATION:  /label= Basic II
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  657..723
;      OTHER INFORMATION:  /label= Mixed Charge
US-08-258-442-11

  Query Match             11.8%;  Score 188;  DB 1;  Length 723;
  Best Local Similarity   23.5%;  Pred. No. 4.2e-08;
  Matches   76;  Conservative   53;  Mismatches  110;  Indels   84;  Gaps   14;

Qy      19 EEDSLVHISQASLDCTVKSGESVVLSVTV-GGAKLVIGTLSQDKFPQISFDLVFDKEFEL 77
           || | |: :::       |  |    ||:  |   :  :: :::: :: || :  |:  :
Db     372 EEISSVNFARSG-----GSTRSFDFEVTLKNGTVHIFSSIEKEEYAKL-FDYITQKKLHV 425

Qy      78 SHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTAVTANGNAGAAVVKADTK 137
           |: |   :   |||       ||  ||:|: |:|                 | :||: :
Db     426 SNMGKDKS----GYKDV-----DFGDSDNENEPDAY--------------LARLKAEAR 461

Qy     138 PKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMDVD---EDDSDDDE------- 187
           |        |:  : | ::|| |||    :::|:   : |   | |||||
Db     462 EK--------EEDDDDGDSDEESTDEDFKPNENESDVAEEYDSNVESDSDDDSDASGGGG 513

Qy     188 EEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKK--AKPAAAPAS-------TPQKT 238
           : |   :::|:  :| |   ||   :  :| |   ||   ||  |  :       | :
Db     514 DSDGAKKKKEKKSEKKEKKEKKHKEKERTKKPSKKKKDSGKPKRATTAFMLWLNDTRESI 573

Qy     239 EKKKGGHTATPHPAKKGGK----------------------SPVNANQSPKSGGQS---S 273
           :::  |   |  |||||:                        |  |::|| |
Db     574 KRENPGIKVT-EIAKKGGEMWKELKDKSKWEDAAAKDKQRYHDEMRNYKPEAGGDSDNEK 632

Qy     274 GGNNNKKPFNSGKQFGGSNNKGS 296
           || ::||          :|  ||
```

```
Db    633 GGKSSKKRKTEPSPSKKANTSGS 655


RESULT 10
US-08-328-809-6
; Sequence 6, Application US/08328809
; Patent No. 5705334
;  GENERAL INFORMATION:
;    APPLICANT:  Lippard, Stephen J.
;    APPLICANT:  Essigmann, John M.
;    APPLICANT:  Donahue, Brian A.
;    APPLICANT:  Toney, Jeffrey H.
;    APPLICANT:  Bruhn, Suzanne L.
;    APPLICANT:  Pil, Pieter M.
;    APPLICANT:  Brown, Steven
;    APPLICANT:  Kellett, Patti
;    TITLE OF INVENTION:  Uses For DNA Structure-Specific
;    TITLE OF INVENTION:  Recognition Proteins
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patent Administrator, Testa, Hurwitz & Thibeault
;      STREET:  53 State Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/328,809
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fenton, Gillian M.
;      REGISTRATION NUMBER:  36,508
;      REFERENCE/DOCKET NUMBER:  MIT-023 (5473/24)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-248-7000
;      TELEFAX:  617-248-7100
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  723 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Drosophila melanogaster
;    IMMEDIATE SOURCE:
;      CLONE:  Drosophila SSRP (predicted)
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  458..507
;      OTHER INFORMATION:  /label= Acidic
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  518..547
;      OTHER INFORMATION:  /label= Basic I
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  547..620
;      OTHER INFORMATION:  /label= HMG
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  632..649
;      OTHER INFORMATION:  /label= Basic II
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  657..723
;      OTHER INFORMATION:  /label= Mixed Charge
US-08-328-809-6


  Query Match             11.8%;  Score 188;  DB 1;  Length 723;
  Best Local Similarity   23.5%;  Pred. No. 4.2e-08;
  Matches   76;  Conservative   53;  Mismatches  110;  Indels   84;  Gaps   14;

Qy      19 EEDSLVHISQASLDCTVKSGESVVLSVTV-GGAKLVIGTLSQDKFPQISFDLVFDKEFEL 77
           || | |: :::       |  \    ||:  |   :  :: :::: :: || :  |:  :
Db     372 EEISSVNFARSG-----GSTRSFDFEVTLKNGTVHIFSSIEKEEYAKL-FDYITQKKLHV 425

Qy      78 SHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTAVTANGNAGAAVVKADTK 137
           |: |   :    |||       ||  ||:|: |:|                 | :||: :
Db     426 SNMGKDKS----GYKDV-----DFGDSDNENEPDAY--------------LARLKAEAR 461

Qy     138 PKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMDVD---EDDSDDDE------- 187
           |        |:  : | ::|| |||    :::|:   : |   | |||||
Db     462 EK--------EEDDDDGDSDEESTDEDFKPNENESDVAEEYDSNVESDSDDDSDASGGGG 513

Qy     188 EEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKK--AKPAAAPAS-------TPQKT 238
           : |   :::|:  :| |   ||   :  :| |   ||   ||  | :       | :
Db     514 DSDGAKKKKEKKSEKKEKKEKKHKEKERTKKPSKKKKDSGKPKRATTAFMLWLNDTRESI 573

Qy     239 EKKKGGHTATPHPAKKGGK----------------------SPVNANQSPKSGGQS---S 273
           :::  |   |  |||||:                          |  |::|| |
Db     574 KRENPGIKVT-EIAKKGGEMWKELKDKSKWEDAAAKDKQRYHDEMRNYKPEAGGDSDNEK 632

Qy     274 GGNNNKKPFNSGKQFGGSNNKGS 296
           || ::||          :|  ||
Db     633 GGKSSKKRKTEPSPSKKANTSGS 655


RESULT 11
PCT-US92-11107-11
; Sequence 11, Application PC/TUS9211107
;  GENERAL INFORMATION:
;    APPLICANT:  Donahue, Brian A.
;    APPLICANT:  Toney, Jeffrey H.
;    APPLICANT:  Bruhn, Suzanne L.
;    APPLICANT:  Pil, Pieter M.
;    APPLICANT:  Brown, Steven
;    APPLICANT:  Kellett, Patti
;    APPLICANT:  Essigmann, John M.
;    APPLICANT:  Lippard, Stephen J.
;    TITLE OF INVENTION:  DNA Structure Specific Recognition
;    TITLE OF INVENTION:  Protein and Uses Therefor
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  2 Militia Drive
;      CITY:  Lexington
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/11107
;      FILING DATE:  19921218
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/539,906
;      FILING DATE:  18-JUN-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  MIT-4787AAA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  11:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  723 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Drosophila melanogaster
;    IMMEDIATE SOURCE:
;      CLONE:  Drosophila SSRP (predicted)
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  458..507
;      OTHER INFORMATION:  /label= Acidic
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  518..547
;      OTHER INFORMATION:  /label= Basic I
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  547..620
;      OTHER INFORMATION:  /label= HMG-box
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  632..649
;      OTHER INFORMATION:  /label= Basic II
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  657..723
;      OTHER INFORMATION:  /label= Mixed Charge
PCT-US92-11107-11


  Query Match             11.8%;  Score 188;  DB 5;  Length 723;
  Best Local Similarity   23.5%;  Pred. No. 4.2e-08;
  Matches   76;  Conservative   53;  Mismatches  110;  Indels   84;  Gaps   14;

Qy      19 EEDSLVHISQASLDCTVKSGESVVLSVTV-GGAKLVIGTLSQDKFPQISFDLVFDKEFEL 77
           || | |: :::       | |    ||:  |   :  :: :::: :: || :  |:  :
Db     372 EEISSVNFARSG-----GSTRSFDFEVTLKNGTVHIFSSIEKEEYAKL-FDYITQKKLHV 425

Qy      78 SHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTAVTANGNAGAAVVKADTK 137
           |: |   :    |||       ||  ||:|: |:|               | :||: :
Db     426 SNMGKDKS----GYKDV-----DFGDSDNENEPDAY--------------LARLKAEAR 461

Qy     138 PKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMDVD---EDDSDDDE------- 187
           |         |:  : | ::|| |||    :::|:   : |   | |||||
Db     462 EK--------EEDDDDGDSDEESTDEDFKPNENESDVAEEYDSNVESDSDDDSDASGGGG 513

Qy     188 EEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKK--AKPAAAPAS-------TPQKT 238
           : |   :::|:  :| |   ||   :  :| |   ||   ||  | :        | :
Db     514 DSDGAKKKKEKKSEKKEKKEKKHKEKERTKKPSKKKKDSGKPKRATTAFMLWLNDTRESI 573

Qy     239 EKKKGGHTATPHPAKKGGK----------------------SPVNANQSPKSGGQS---S 273
           :::  |   |   |||||:                      |  |::|| |
Db     574 KRENPGIKVT-EIAKKGGEMWKELKDKSKWEDAAAKDKQRYHDEMRNYKPEAGGDSDNEK 632

Qy     274 GGNNNKKPFNSGKQFGGSNNKGS 296
           || ::||          :|  ||
Db     633 GGKSSKKRKTEPSPSKKANTSGS 655


RESULT  12
US-08-728-323A-2
; Sequence 2, Application US/08728323A
; Patent No. 5948676
;  GENERAL INFORMATION:
;    APPLICANT:  Chang, Yuan
;    APPLICANT:  Bohenzky, Roy A.
;    APPLICANT:  Russo, James J.
;    APPLICANT:  Edelman, Isidore S.
;    APPLICANT:  Moore, Patrick S.
;    TITLE OF INVENTION:  Immediate Early Protein From Kaposi's
;    TITLE OF INVENTION:  Sarcoma-Associated Herpesvirus, DNA
;    TITLE OF INVENTION:  Encoding Same And Uses Thereof
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/728,323A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  0575/52268/JPW/MSC/SKS
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-278-0400
;      TELEFAX:  212-391-0525
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH: 1162 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-728-323A-2


  Query Match             11.3%;  Score 180;  DB 2;  Length 1162;
  Best Local Similarity   27.8%;  Pred. No. 3.7e-07;
  Matches   49;  Conservative   18;  Mismatches   55;  Indels   54;  Gaps    3;

Qy     103 SSDDEDVPEAVPAPAPTAV----------------------------------TANGNAG 128
           || | | |   |  :| ::                                  : |  ||
Db     248 SSSDGDTPPRQPPTSPISIGSSSPSEGSWGDDTAMLVLLAEIAEEASKNEKECSENNQAG 307

Qy     129 -----------AAVVKADTKPKAKPAEVKPAEEKPESDEE---------DESDDEDESEE 168
                     : | | |   |    | : ||  | |||           || |||:: ||
Db     308 EDNGDNEISKESQVDKDDNDNKDDEEEQETDEEDEEDDEEDDEEDDEEDDEEDDEEDDEE 367

Qy     169 DDDSEKGMDVDEDDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKK 224
           ||: :   | :||| :|||||| |::|||:  :: |  :    ||          ||
Db     368 DDEEDDEEDDEEDDEEDDEEEDEEEDEEEDEEEEDEEDDDDEDNEDEEDDEEEDKK 423


RESULT  13
US-08-375-300-4
; Sequence 4, Application US/08375300
; Patent No. 5679566
;  GENERAL INFORMATION:
;    APPLICANT:  Feng, He
;    APPLICANT:  Jacobson, Allan S.
;    TITLE OF INVENTION:  HETEROLOGOUS POLYPEPTIDE PRODUCTION IN
;    TITLE OF INVENTION:  THE ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street Suite 3100
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/375,300
;      FILING DATE:  20-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fasse, J. P.
;      REGISTRATION NUMBER:  32,983
;      REFERENCE/DOCKET NUMBER:  04020/046001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)542-5070
;      TELEFAX:  (617)542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  764 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-375-300-4


  Query Match             10.7%;  Score 170;  DB 1;  Length 764;
  Best Local Similarity   25.3%;  Pred. No. 1.5e-06;
  Matches   64;  Conservative   32;  Mismatches   71;  Indels   86;  Gaps   11;

Qy      81 GTKANVHFIGYKS----PNIEQDDFTSSDDED-------VPEAVPAPAPTAVTANGNAGA 129
           |: |::|  | :|      :| :||     |: |          |       |   :| |
Db     498 GSSASIH-NGKESAVPIESITEDDEDEDDENDDGVDLLGEDEDAEISTPNTESAPG---- 552

Qy     130 AVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMDVDEDDSDDDEEE 189
               | :||  |      |  : :::|| ||:|: ::||| |:|   |||| :||::|
Db     553 -------KHQAKQDE----SEDEDDEDDDEDDDDDDDDDDDDGEEG---DEDDDEDDDDE 598

Qy     190 DSEDEEE----------------------------------EETPKKPEPINKK------ 209
           | :||||                                  ||  |  | : ::
Db     599 DDDDEEEEDSDSDLEYGGDLDADRDIEMKRMYEEYERKLKDEEERKAEEELERQFQKMMQ 658

Qy     210 -----RPNESV--SKTPVSGKKA--------KPAAAPASTPQKTEK-KKGGHTATPHPAK 253
                | :| |  || ||  |          | :  |:|: :  |:  |    |     |
Db     659 ESIDARKSEKVVASKIPVISKPVSVQKPLLLKKSEEPSSSKETYEELSKPKKIAFTFLTK 718

Qy     254 KGGKSPVNANQSP 266
           | |:     | |
Db     719 SGKKTQSRILQLP 731


RESULT  14
US-09-177-431-4
; Sequence 4, Application US/09177431
; Patent No. 6071700
;  GENERAL INFORMATION:
;    APPLICANT:  He, Feng
;    APPLICANT:  Jacobson, Allan S.
;    TITLE OF INVENTION:  HETEROLOGOUS POLYPEPTIDE PRODUCTION IN THE
;    TITLE OF INVENTION:  ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/177,431
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/955,472
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fasse, J. Peter
;      REGISTRATION NUMBER:  32,983
;      REFERENCE/DOCKET NUMBER:  07917/050001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-9806
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  764 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-177-431-4


  Query Match             10.7%;  Score 170;  DB 3;  Length 764;
  Best Local Similarity   25.3%;  Pred. No. 1.5e-06;
  Matches   64;  Conservative   32;  Mismatches   71;  Indels   86;  Gaps   11;

Qy      81 GTKANVHFIGYKS----PNIEQDDFTSSDDED-------VPEAVPAPAPTAVTANGNAGA 129
           |: |::|  | :|      :| :||     |: |          |       |   :| |
Db     498 GSSASIH-NGKESAVPIESITEDDEDEDDENDDGVDLLGEDEDAEISTPNTESAPG---- 552

Qy     130 AVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMDVDEDDSDDDEEE 189
               | :||  |      |  : :::|| ||:|: ::||| |:|   |||| :||::|
Db     553 -------KHQAKQDE----SEDEDDEDDDEDDDDDDDDDDDDGEEG---DEDDDEDDDDE 598

Qy     190 DSEDEEE----------------------------------EETPKKPEPINKK------ 209
           | :||||                                  ||  |  | : ::
Db     599 DDDDEEEEDSDSDLEYGGDLDADRDIEMKRMYEEYERKLKDEEERKAEEELERQFQKMMQ 658

Qy     210 -----RPNESV--SKTPVSGKKA--------KPAAAPASTPQKTEK-KKGGHTATPHPAK 253
                | :| |  || ||  |          | :  |:|: :  |:  |    |     |
Db     659 ESIDARKSEKVVASKIPVISKPVSVQKPLLLKKSEEPSSSKETYEELSKPKKIAFTFLTK 718

Qy     254 KGGKSPVNANQSP 266
           | |:     | |
Db     719 SGKKTQSRILQLP 731


RESULT  15
PCT-US95-16930-4
; Sequence 4, Application PC/TUS9516930
;  GENERAL INFORMATION:
;    APPLICANT:  UNIVERSITY OF MASSACHUSETTS MEDICAL SCHOOL
;    TITLE OF INVENTION:  HETEROLOGOUS POLYPEPTIDE
;    TITLE OF INVENTION:  PRODUCTION IN THE ABSENCE OF
;    TITLE OF INVENTION:  NONSENSE-MEDIATED mRNA DECAY
;    TITLE OF INVENTION:  FUNCTION
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street Suite 3100
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/16930
;      FILING DATE:  27-DEC-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/375,300
;      FILING DATE:  20-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fasse, J. Peter
```

```
;      REGISTRATION NUMBER:  32,983
;      REFERENCE/DOCKET NUMBER:  04020/046WO1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)542-5070
;      TELEFAX:  (617)542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  764 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US95-16930-4


  Query Match            10.7%;  Score 170;  DB 5;  Length 764;
  Best Local Similarity  25.3%;  Pred. No. 1.5e-06;
  Matches   64;  Conservative   32;  Mismatches   71;  Indels   86;  Gaps   11;

Qy      81 GTKANVHFIGYKS----PNIEQDDFTSSDDED-------VPEAVPAPAPTAVTANGNAGA 129
           |: |::|  | :|     :| :||    |: |         |      |   :| |
Db     498 GSSASIH-NGKESAVPIESITEDDEDEDDENDDGVDLLGEDEDAEISTPNTESAPG---- 552

Qy     130 AVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMDVDEDDSDDDEEE 189
                 | :||  |     |  : :::|| ||:|: ::||| |:|   |||| :||::|
Db     553 -------KHQAKQDE----SEDEDDEDDDEDDDDDDDDDDDGEEG---DEDDDEDDDDE 598

Qy     190 DSEDEEE---------------------------------EETPKKPEPINKK------ 209
           | :||||                                ||  |  | : ::
Db     599 DDDDEEEEDSDSDLEYGGDLDADRDIEMKRMYEEYERKLKDEEERKAEEELERQFQKMMQ 658

Qy     210 -----RPNESV--SKTPVSGKKA--------KPAAAPASTPQKTEK-KKGGHTATPHPAK 253
                | :| |  || ||  |          | :  |:|: :  |:  |     |     |
Db     659 ESIDARKSEKVVASKIPVISKPVSVQKPLLLKKSEEPSSSKETYEELSKPKKIAFTFLTK 718

Qy     254 KGGKSPVNANQSP 266
           | |:    | |
Db     719 SGKKTQSRILQLP 731


Search completed: May  2, 2002, 22:08:42
Job time: 6118 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: May 2, 2002, 22:07:30 ; Search time 205.36 Seconds
(without alignments)
142.712 Million cell updates/sec

Title: US-09-645-337-8
Perfect score: 1589
Sequence: 1 MEFWGVAVTPKNATKVTPEE.....KQFGGSNNKGSNKGKGKGRA 305

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1589	100.0	305	2	T52287	probable histone d
2	543.5	34.2	307	2	T04141	histone deacetylas
3	450.5	28.4	296	2	T48401	histone deacetylas
4	311	19.6	412	2	A55320	immunophilin FKBP4
5	243	15.3	357	2	JC4090	FK506-binding 39k
6	230	14.5	798	2	T33022	hypothetical prote
7	215.5	13.6	694	1	DNCHNL	nucleolin - chicke
8	211.5	13.3	712	2	JH0148	nucleolin - rat
9	207.5	13.1	611	2	T06458	nucleolin homolog
10	206.5	13.0	707	1	DNMS	nucleolin - mouse
11	206	13.0	713	2	A27441	nucleolin - Chines
12	205	12.9	707	2	A35804	nucleolin - human
13	200.5	12.6	281	2	JC4295	heat-shock protein
14	196.5	12.4	635	2	T09648	nucleolin homolog
15	192.5	12.1	971	2	T19431	hypothetical prote
16	188	11.8	723	2	S33688	hypothetical prote
17	188	11.8	734	2	B42680	nucleolus-cytoplas
18	186	11.7	723	2	A48217	single-strand DNA/
19	185	11.6	1151	2	T24541	hypothetical prote
20	183.5	11.5	294	1	DNCHFM	nucleophosmin - ch
21	182.5	11.5	872	2	T18861	probable peptide-a
22	182.5	11.5	1187	2	T46637	transcription fact
23	180.5	11.4	374	2	C88734	protein F32E10.6 [
24	180	11.3	1300	2	T03166	probable immediate
25	178.5	11.2	517	2	S20248	transcription fact
26	177.5	11.2	885	2	G71608	ATP-dept. acyl-CoA
27	177.5	11.2	1473	2	T13855	suppressor of sabl
28	177	11.1	644	2	S55395	neurofilament prot
29	177	11.1	651	2	S18874	nucleolin - Africa
30	177	11.1	705	2	S32644	nucleolin - Africa
31	176	11.1	896	2	T24169	hypothetical prote
32	176	11.1	906	2	T24166	hypothetical prote
33	176	11.1	1182	2	T30189	myelin transcripti
34	175	11.0	524	2	S35551	transcription fact
35	174.5	11.0	763	2	T08929	hypothetical prote
36	174	11.0	211	2	T25911	hypothetical prote
37	174	11.0	1188	2	T46608	zinc finger protei
38	173.5	10.9	361	2	T43536	peptidyl prolyl ci
39	172	10.8	292	2	I52858	nucleolar protein
40	172	10.8	557	2	A96527	probable nuM1 prot
41	172	10.8	678	2	A54514	glutamic acid-rich
42	171	10.8	210	2	T28771	hypothetical prote
43	170.5	10.7	3498	2	T22330	hypothetical prote
44	170	10.7	1089	2	S48244	NMD2 protein - yea
45	170	10.7	1110	2	I51116	NF-180 - sea lampr

ALIGNMENTS

RESULT 1
T52287
probable histone deacetylase (EC 3.5.1.-) [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
C;Accession: T52287
R;Dangl, M.; Haas, H.; Loidl, P.
submitted to the EMBL Data Library, January 1998
A;Description: Arabidopsis thaliana complete cDNA-sequence homologous to Zea mays HD2
A;Reference number: Z26016
A;Accession: T52287
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-305 <DAN>
A;Cross-references: EMBL:AF044914; PIDN:AAC02539.1
A;Experimental source: cultivar Columbia
C;Genetics:
A;Gene: HD2
C;Keywords: hydrolase

Query Match 100.0%; Score 1589; DB 2; Length 305;
Best Local Similarity 100.0%; Pred. No. 6.6e-78;
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQD 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQD 60

Qy      61 KFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 KFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTA 120

Qy     121 VTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMDVDE 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 VTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMDVDE 180

Qy     181 DDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTEK 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 DDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTEK 240

Qy     241 KKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGGNNNKKPFNSGKQFGGSNNKGSNKGK 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 KKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGGNNNKKPFNSGKQFGGSNNKGSNKGK 300

Qy     301 GKGRA 305
           |||||
Db     301 GKGRA 305
```

RESULT 2
T04141

```
histone deacetylase (EC 3.5.1.-) HD2-p39, nucleolar - maize
C;Species: Zea mays (maize)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C;Accession: T04141
R;Lusser, A.; Brosch, G.; Loidl, A.; Haas, H.; Loidl, P.
Science 277, 88-91, 1997
A;Title: Identification of maize histone deacetylase HD2 as an acidic nucleolar phosphop
A;Reference number: Z15237; MUID:97349336
A;Accession: T04141
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-307 <LUS>
A;Cross-references: EMBL:U82815; NID:g2257755; PIDN:AAB63262.1; PID:g2257756
A;Experimental source: strain Cuzco 251
C;Keywords: hydrolase

  Query Match             34.2%;  Score 543.5;  DB 2;  Length 307;
  Best Local Similarity   44.8%;  Pred. No. 2.4e-22;
  Matches  130;  Conservative   45;  Mismatches   96;  Indels   19;  Gaps   11;

Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQD 60
           |||||: | | :  |  |    ::|:|||:|  : || :: :: | :   || ||||| |
Db       1 MEFWGLEVKPGSTVKCEPGYGFVLHLSQAALGESKKS-DNALMYVKIDDQKLAIGTLSVD 59

Qy      61 KFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTA 120
           | | | |||:||||||||||:    :| | |||     ::|    | ||  | :  |
Db      60 KNPHIQFDLIFDKEFELSHTSKTTSVFFTGYKVEQPFEEDEMDLDSEDEDEELNVP---V 116

Qy     121 VTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEED--DDSEKGMDV 178
           |  || |      |:  |  | |::  |  :| : |:: : |: |:|:||  |||::|:
Db     117 VKENGKADEKKQKSQEKAVAAPSKSSPDSKKSKDDDDSDEDETDDSDEDETDDSDEGLSS 176

Qy     179 DEDDSDDDEEEEDSEDEEEEE--TPKKPEPINKKRPNE-SVSKTPVSGKKAKPAAAPASTP 235
           :| | |  :|:|: |:|||:  ||||||  : |||| | || |||:| ||||  | |:|
Db     177 EEGDDDSSDEDDTSDDEEEDTPTPKKPE-VGKKRPAESSVLKTPLSDKKAK-VATPSS-- 232

Qy     236 QKTEKKKGG--HTATPHPAKKGGKSPVNANQSPKS--GGQSSGGNNNKKP 281
           |||  |||   | |||||||  ||: || ::| ||      |||:   ||
Db     233 QKTGGKKGAAVHVATPHPAK--GKTIVNNDKSVKSPKSAPKSGGSVPCKP 280


RESULT   3
T48401
histone deacetylase-like protein - Arabidopsis thaliana
N;Alternate names: protein F17C15.160
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C;Accession: T48401
R;Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.;
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24492
A;Accession: T48401
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-296 <BEV>
A;Cross-references: EMBL:AL162506
A;Experimental source: cultivar Columbia; BAC clone F17C15
C;Genetics:
A;Map position: 5
A;Introns: 5/1; 28/3; 93/3; 116/3; 164/3; 195/3; 227/1
A;Note: F17C15.160

  Query Match             28.4%;  Score 450.5;  DB 2;  Length 296;
  Best Local Similarity   38.3%;  Pred. No. 2e-17;
  Matches  118;  Conservative   39;  Mismatches   94;  Indels   57;  Gaps   12;

Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGES-------VVLSVTVGGAKLV 53
           |||||| |       : |  | |||||| :|      |||       : | ||||  ||:
Db       1 MEFWGVEVKNGKPLHLDPGLDRLVHISQVAL------GESKNNVTEPIQLYVTVGSDKLL 54

Qy      54 IGTLSQDKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDED----- 108
           ||||| :||||:| ::| :: | |||:    :| | |||  ::  |   ::  :
Db      55 IGTLSHEKFPQLSTEIVLERNFALSHTWKNGSVFFSGYKPEDLIDDQLEAAGFKGKWGLL 114

Qy     109 VPEAVPAPAPTAVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEE 168
           | | |  |    |          :  | | |         ::  :  ||| |||:|
Db     115 YPPAAPKSAAKQVNFQ------LPNEDVKAK--------QDDDADGSEEDSSDDDDSENS 160

Qy     169 DDDSEKGM----DVDEDDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKK 224
           |: |: :    | :|||| ||||:||    |||||||||   |:    : || | | ||
Db     161 GDEEEEKVTAESDSEEDDSSDDEEDDS---SEEETPKKPEEPKKRSAEPNSSKNPASNKK 217

Qy     225 AKPAAAPASTPQKTEKKKGG-HTATPHPAKKGGKSPVNANQSPKSGGQSSGGNNNKK--P 281
           ||       |||||: ||   | |||||:|: ||:          ||| |:|  : ::  |
Db     218 AK-----FVTPQKTDSKKPHVHVATPHPSKQAGKN---------SGGGSTGETSKQQQTP 263

Qy     282 FNSGKQFG 289
           ::|  ||
Db     264 KSAG-AFG 270


RESULT   4
A55320
immunophilin FKBP46 - fall armyworm
C;Species: Spodoptera frugiperda (fall armyworm)
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 02-Sep-2000
C;Accession: A55320
R;Alnemri, E.S.; Fernandes-Alnemri, T.; Pomerenke, K.; Robertson, N.M.; Dudley, K.; D
J. Biol. Chem. 269, 30828-30834, 1994
A;Title: FKBP46, a novel Sf9 insect cell nuclear immunophilin that forms a protein-ki
A;Reference number: A55320; MUID:95074110
A;Accession: A55320
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-412 <ALN>
A;Cross-references: GB:U15030; NID:g595844; PID:g595845
C;Superfamily: yeast peptidylprolyl isomerase FPR3; BKBP-type peptidylprolyl isomeras
C;Keywords: nucleus; phosphoprotein
F;324-371/Domain: BKBP-type peptidylprolyl isomerase homology <PPI>

  Query Match             19.6%;  Score 311;  DB 2;  Length 412;
  Best Local Similarity   31.0%;  Pred. No. 7.1e-10;
  Matches  100;  Conservative   39;  Mismatches  116;  Indels   68;  Gaps   14;

Qy       3 FWGVAVTP-KNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQDK 61
           |||: : | |  |:|    :   ||||::|  :    :   : | | |   :: || : |
Db       2 FWGLIMEPNKRYTQVV---EKPFHISQAAMDISTGDNDPCQVMVVVDGKNFLVCTLQKGK 58

Qy      62 FPQISFDLVFDKEFELSH-SGTKANVHFIGYKSPNIEQD--DFTSSDDEDVPEAVPAPAP 118
             |:  || |    :|  :  | |||  ||  |  |:|  |   :::|:  |  |   |
Db      59 IIQVPLDLYFKSGDSVSFLTNGKCNVHLTGYLDPEFEEDLEDEEEAEEEEEEEAPPLVP 118

Qy     119 T---AVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEED----------ESDDED 164
                 | |   |  ||    ||       ||   : |:||           ::|: |
Db     119 AKNKRKLENANDATANKKAKPDKKAGKNSAPAAESDSDDDEDQLQKFLDGEDIDTDEND 178

Qy     165 ESEEDDDSEKGMDVDEDDSDDDEE----EDSEDEEEEETPKKPEPINKKRP--------- 211
           || : : | :| | ||:| |:|||    :| :|||||| |||     ||:|
Db     179 ESFKMNTSAEGDDSDEEDDDEDEEDEEDDDEDDEEEEEAPKK----KKKQPAAEQDSTLD 234

Qy     212 --NESV--SKTPVSGK----------KAKPAAAPASTP------QKTEKKKGGHTATPH 250
              |||  ||   | |          | :|       |      || ||||
Db     235 TSKESVDMSKLSKSQKRRLKKKLQQQAKQQPQVNGVDKPKKEEPQQKAEKKK-------- 286

Qy     251 PAKKGGKSPVNANQSPK-SGGQS 272
           |  |  ::||   :  : :|| |
Db     287 PEAKKEEAPVEKKEKKQIAGGVS 309


RESULT   5
JC4090
FK506-binding 39k protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
```

```
C;Date: 23-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 21-Jul-2000
C;Accession: JC4090
R;Theopold, U.; Zotto, L.D.; Hultmark, D.
Gene 156, 247-251, 1995
A;Title: FKBP39, a Drosophila member of a family of proteins that bind the immunosuppres
A;Reference number: JC4090; MUID:95278752
A;Accession: JC4090
A;Molecule type: mRNA
A;Residues: 1-357 <THE>
A;Cross-references: EMBL:Z46894; NID:g600423; PIDN:CAA86996.1; PID:g600424
C;Genetics:
A;Gene: fkbp39
A;Cross-references: FlyBase:FBgn0013269
C;Superfamily: yeast peptidylprolyl isomerase FPR3; BKBP-type peptidylprolyl isomerase h
F;89-99/Region: acidic
F;252-357/Region: FK506 binding #status predicted
F;269-316/Domain: BKBP-type peptidylprolyl isomerase homology <PPI>


  Query Match             15.3%;  Score 243;  DB 2;  Length 357;
  Best Local Similarity   26.9%;  Pred. No. 2.5e-06;
  Matches   95;  Conservative   49;  Mismatches  109;  Indels  100;  Gaps   18;

Qy       3 FWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQDKF 62
           |||: : |:   | :       |||  :||     |:   | :     : :: |:::
Db       4 FWGLNMKPER--KYSQTIIKSFHISGVALD----KGQEAKLYLAAEKQEYIVATVTK-AI 56

Qy      63 PQISFDLVFDK-EFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTAV 121
           ||:: || | | :  : ::   |:|  :||        | | ||||  :          :
Db      57 PQVALDLNFSKGDRIMFYTAGDASVSLLGY------LHDIDSEDDEDDDQ---------M 101

Qy     122 TANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDES-------------E 167
           |      :  :|   |        :  |:: || |||| | :|:|               |
Db     102 TIENLLNSKAIKNSKKS-------EDDEDENESGEEDEEDTDDDSQIIEEYESFLENGEE 154

Qy     168 EDDDSEKGMDVDED--DSDDDEEEDSEDEE-EEETPKKPEPINKKRPNESVSKTPVSGKK 224
           ||||     |||||  :| :::|:||:| | ||| ||:   : |  |  |  |   |||:
Db     155 EDDD-----DVDEDNEESGEEDEQDSDDSEAEEEQPKR--KVAKLSPGASAKK---SGKE 204

Qy     225 AKPAAAPASTPQKTEKKKGGHTATPHPAKKGGKSPVNANQSPKS---------------- 268
               |      |: :||:      |   |:  |:   | | | |
Db     205 QNGVAKKEEAKQQQKKKE-----KPEAKKEQPKAKEPAKQQPASKDPRTITGGVKIVDQV 259

Qy     269 ---GGQSSGG-----------NNNKKPFNS---GKQF-----GGSNNKGSNKG 299
               | ::  |           :| | |:|   || |     ||   || : |
Db     260 VGKGEEAKQGKRVSVYYIGRLQSNNKTFDSLLKGKPFKFALGGGEVIKGWDVG 312


RESULT   6
T33022
hypothetical protein K07H8.10 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T33022
R;Fulton, B.; Hawkins, J.; Gattung, S.; Wohldmann, P.; Elliott, G.
submitted to the EMBL Data Library, February 1998
A;Description: The sequence of C. elegans cosmid K07H8.
A;Reference number: Z21264
A;Accession: T33022
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-798 <FUL>
A;Cross-references: EMBL:AF047659; PIDN:AAC04430.1; GSPDB:GN00022; CESP:K07H8.10
A;Experimental source: strain Bristol N2; clone K07H8
C;Genetics:
A;Gene: CESP:K07H8.10
A;Map position: 4
A;Introns: 205/1; 308/3; 630/1; 773/3


  Query Match             14.5%;  Score 230;  DB 2;  Length 798;
  Best Local Similarity   31.7%;  Pred. No. 2.7e-05;
  Matches   72;  Conservative   32;  Mismatches   83;  Indels   40;  Gaps    9;

Qy      48 GGAKLVIGTLSQDKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDE 107
           ||:::       ||  |   |   |:| |:          |     | :| |  :  |:|
Db     181 GGSRVGF----SDKRKQYDSDGDDDEEEEVPKKKLATGTPFAKQTKP-VEDD--SGEDEE 233

Qy     108 DVPEAVPAPAPTAVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESE 167
           |  |:   | |                  |  || |    ||   :||:::| |||:| |
Db     234 DEEESDEEPQP----------------PKKTAKSAIATAAE---DSDDDEEDDDEEEEEE 273

Qy     168 EDDDS---EKGMDVDEDDSDDDEEE----DSEDEEEEETPKKPEPINKKRPNESVSKTPV 220
           |||:   :| : |:||: |:::||    | |||||||  :: ||: |    :|  |:
Db     274 SDDDAPQVKKSVAVEEDEEDEEDEEEIEMDEEDEEEEEDEEETEPVAKTPAVKSSLKSVD 333

Qy     221 SGKKAKPAAAPASTPQKTEKKKGGHTATPHPAKKGGKSPVNANQSPK 267
           | |  :   :  :||   :||     |||||||    :   |   ||
Db     334 STKGKQVNLSKVATPTTVDKKA---PATPHPAKTSKNA---AGDVPK 374


RESULT   7
DNCHNL
nucleolin - chicken
N;Alternate names: nucleolar protein C23
C;Species: Gallus gallus (chicken)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 22-Jun-1999
C;Accession: S08414; S10766; A32725; I50397; B30099
R;Maridor, G.; Nigg, E.A.
Nucleic Acids Res. 18, 1286, 1990
A;Title: cDNA sequences of chicken nucleolin/C23 and NO38/B23, two major nucleolar pr
A;Reference number: S08414; MUID:90206792
A;Accession: S08414
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-694 <MAR>
A;Cross-references: EMBL:X17199; NID:g63710; PIDN:CAA35060.1; PID:g63711
R;Maridor, G.; Krek, W.; Nigg, E.A.
Biochim. Biophys. Acta 1049, 126-133, 1990
A;Title: Structure and developmental expression of chicken nucleolin and NO38: coordi
A;Reference number: S10766; MUID:90304215
A;Accession: S10766
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-694 <MAR2>
R;Peter, M.; Nakagawa, J.; Doree, M.; Labbe, J.C.; Nigg, E.A.
Cell 60, 791-801, 1990
A;Title: Identification of major nucleolar proteins as candidate mitotic substrates o
A;Reference number: A32725; MUID:90182668
A;Accession: A32725
A;Molecule type: protein
A;Residues: 56-62;63-109;197-214 <PET>
R;Borer, R.A.; Lehner, C.F.; Eppenberger, H.M.; Nigg, E.A.
Cell 56, 379-390, 1989
A;Title: Major Nucleolar Proteins Shuttle between Nucleus and Cytoplasm.
A;Reference number: I50397; MUID:89119560
A;Accession: I50397
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 407-418,'R',420-519,'T',521-694 <BOR>
A;Cross-references: GB:M21791; NID:g212411; PIDN:AAA48983.1; PID:g212412
C;Comment: Phosphorylation of this protein by cdc2 kinase may contribute to the reorg
C;Superfamily: nucleolin; ribonucleoprotein repeat homology
C;Keywords: DNA binding; duplication; nucleolus; nucleus; phosphoprotein; RNA binding
F;1-247/Domain: acidic <ACI>
F;54-91/Region: 7-residue repeats (T-P-A-K-K-A-[A/V])
F;254-262/Region: nuclear location signal
F;267-275/Region: nuclear location signal
F;282-347/Domain: ribonucleoprotein repeat homology <RRM1>
F;283-361/Domain: RNA binding #status predicted <RNA1>
F;372-435/Domain: ribonucleoprotein repeat homology <RRM2>
F;373-448/Domain: RNA binding #status predicted <RNA2>
F;462-525/Domain: ribonucleoprotein repeat homology <RRM3>
F;463-538/Domain: RNA binding #status predicted <RNA3>
F;554-618/Domain: ribonucleoprotein repeat homology <RRM4>
F;555-631/Domain: RNA binding #status predicted <RNA4>
```

```
F;632-694/Domain: glycine/arginine-rich <GRR>
F;56,63,70,77,85/Binding site: phosphate (Thr) (covalent) (by cdc2 kinase) #status predi


  Query Match             13.6%;  Score 215.5;  DB 1;  Length 694;
  Best Local Similarity   28.1%;  Pred. No. 0.00014;
  Matches   85;  Conservative   32;  Mismatches  102;  Indels   83;  Gaps   12;

Qy       6 VAVTPKNATK----VTPEEDSLVHISQASLDCTVKSGESVVL--------SVT------- 46
           :| ||||  |      | :     : | |     ||| |::         :||
Db       4 LAKTPKNQMKQKKMAPPPKKVEESEEEESSDLEESSGEEVMVPPKKQQKAAVTPAKKAAT 63

Qy      47 ------------VGGAKLVIGTLSQDKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSP 94
                       |  ||  : |      |          |:  :   | |   :    :|
Db      64 PAKKAATPAKKAVTPAKKAVAT------PAKKAVAPSPKKAAVVGKGAKNGKNAKKEESE 117

Qy      95 NIEQDDFTSSDDEDVPEAVPAPAPTAVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPES 154
             ::||    :|||  |         ||        ||   |   |  ||| | ||  | ||
Db     118 EEDEDDEDDEEDEDEEEESDEEEEPAVPVK---PAAKKSAAAVPAKKPA-VVPA--KQES 171

Qy     155 DEEDESDDEDESEEDDDSE-KGMDV-------------------------DEDDSDDDE 187
           :||:| |||:| ||||:|| : ||                          ||:| |:||
Db     172 EEEEEEDDEEEDEEDDESEDEAMDTTPAPVKKPTPAKATPAKAKAESEDEEDEEDEDEDE 231

Qy     188 EEDSEDEEEEETPKKPEPINK---KRPNESVSKTPVSGKKAKPAAAPAS---------TP 235
           |:: ::||:||  :  :|: :   ||  |  :|:    || |    |||          ||
Db     232 EDEDDEEEDEEESEDEKPVKEAPGKRKKEMANKSAPEAKK-KKTETPASAFSLFVKNLTP 290

Qy     236 QK 237
           |
Db     291 TK 292


RESULT   8
JH0148
nucleolin - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 23-Jul-1999
C;Accession: JH0148; A24088; I63130
R;Bourbon, H.M.; Amalric, F.
Gene 88, 187-196, 1990
A;Title: Nucleolin gene organization in rodents: highly conserved sequences within three
A;Reference number: JH0148; MUID:90269607
A;Accession: JH0148
A;Molecule type: DNA
A;Residues: 1-712 <BOU>
A;Note: the authors translated the initiation codon GTG for residue 1 as Val
R;Lischwe, M.A.; Cook, R.G.; Ahn, Y.S.; Yeoman, L.C.; Busch, H.
Biochemistry 24, 6025-6028, 1985
A;Title: Clustering of glycine and Ng,Ng-dimethylarginine in nucleolar protein C23.
A;Reference number: A24088; MUID:86104094
A;Accession: A24088
A;Molecule type: protein
A;Residues: 651-703 <LIS>
R;Bourbon, H.
Gene 68, 73-84, 1988
A;Title: Sequence and structure of the nucleolin promoter in rodents: Characterization c
A;Reference number: I48118; MUID:89121496
A;Accession: I63130
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 'MV',2-44 <RES>
A;Cross-references: GB:M22090; NID:g205793; PIDN:AAA41733.1; PID:g205794
C;Comment: This protein is the major nucleolar-specific protein in eukaryotic exponentia
C;Genetics:
A;Gene: nuc
A;Start codon: GTG
A;Introns: 5/3; 44/3; 211/1; 274/1; 303/1; 350/2; 392/1; 433/2; 485/1; 526/2; 571/1; 613
C;Superfamily: nucleolin; ribonucleoprotein repeat homology
C;Keywords: DNA binding; nucleus
F;311-376/Domain: ribonucleoprotein repeat homology <RRM1>
F;397-459/Domain: ribonucleoprotein repeat homology <RRM2>
F;489-552/Domain: ribonucleoprotein repeat homology <RRM3>
F;575-639/Domain: ribonucleoprotein repeat homology <RRM4>


  Query Match             13.3%;  Score 211.5;  DB 2;  Length 712;
  Best Local Similarity   29.4%;  Pred. No. 0.00023;
  Matches   57;  Conservative   27;  Mismatches   55;  Indels   55;  Gaps    7;

Qy      97 EQDDFTSSDDEDVPEAVPAPAPTAVTANGNAGAAVVKADTKPKAKPAEVKPA-------E 149
           |:|:  | :|||  :    |               |||        |||:  ||       :
Db     150 EEDEDDSDEDEDEEDEFEPP-------------VVKG-----VKPAKAAPAAPASEDED 190

Qy     150 EKPESDEEDESDDEDESEEDDDSEKGMDV----------------------DEDDSDDD 186
           |: : ||:|: |||:| ||||  |: |::                      :||| ||:
Db     191 EEDDDDEDDDDDDEEEEEEDDSEEEVMEITPAKGKKTPAKVVPVKAKSVAEEEEDDEDDE 250

Qy     187 EEEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQ-KTEKKKGGH 245
           :||: ||||:||     |  :::   | |   |  ||: |         |: | :| :|
Db     251 DEEEDEDEEDEE---DDEDEDEEEEEPVKAAP--GKRKKEMTKQKEAPEAKKQKIEGSE 305

Qy     246 TATPHPAKKGGKSP 259
             ||     |  :|
Db     306 PTTPFNLFIGNLNP 319


RESULT   9
T06458
nucleolin homolog - garden pea
C;Species: Pisum sativum (garden pea)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 08-Oct-1999
C;Accession: T06458
R;Tong, C.G.; Hsieh, H.L.; Blumenthal, S.; Reichler, S.; Balk, J.; Roux, S.J.
submitted to the EMBL Data Library, August 1995
A;Description: Molecular cloning and characterization of a cDNA encoding a nucleolin-
A;Reference number: Z15692
A;Accession: T06458
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-611 <TON>
A;Cross-references: EMBL:L43510; NID:g940287; PIDN:AAA74208.1; PID:g940288
A;Experimental source: cv. Alaska
C;Superfamily: nucleolin; ribonucleoprotein repeat homology


  Query Match             13.1%;  Score 207.5;  DB 2;  Length 611;
  Best Local Similarity   30.9%;  Pred. No. 0.00033;
  Matches   64;  Conservative   24;  Mismatches   54;  Indels   65;  Gaps    8;

Qy      97 EQDDFTSSDDEDVPEAVP-----APAPTAVTANGNAGAAVVKADTK-PKAK--PAEVKP- 147
           | |: :| |::  |:|||     |||  | :::          :| : ||||  || |||
Db     153 ESDEDSSEDEKPAPKAVPAKNGSAPAKKAASSSDEESDEDSSSDEEAPKAKVVPAAVKPA 212

Qy     148 -AEEKPES----------DEED-----------------ESDDEDESEEDDDS------ 172
            |::| ||          ||||                 | |||| |:|:| |
Db     213 AAKKKAESSDSDSSEDSDDEEDTKPTVTAVSKSAAVAKKTEDDDEDSSDEEDSSSDEDNK 272

Qy     173 -----------------EKGMDVDEDDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESV 215
                            |  |:||:| || ||| ||::| | |       |  | :  | :
Db     273 TSKPSNGNKKPVTVSKKEDKMNVDKDSSDSDEESDSDEESESEDEPSKTPQKKTKDVEMI 332

Qy     216 SKTPVSGKKAKPAAAPASTPQKTEKKK 242
                 : | :| | |  :|| :    |
Db     333 D----ADKSSKKAPATPATPSENGGSK 355


RESULT  10
DNMS
nucleolin - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 22-Jun-1999
C;Accession: A29958; A40769; A56240; I84688
R;Bourbon, H.M.; Lapeyre, B.; Amalric, F.
J. Mol. Biol. 200, 627-638, 1988
A;Title: Structure of the mouse nucleolin gene. The complete sequence reveals that ea
```

```
A;Reference number: A29958; MUID:88316930
A;Accession: A29958
A;Molecule type: DNA
A;Residues: 1-707 <BOU>
A;Cross-references: GB:X07699; NID:g53453; PIDN:CAA30538.1; PID:g53454
R;Pasternack, M.S.; Bleier, K.J.; McInerney, T.N.
J. Biol. Chem. 266, 14703-14708, 1991
A;Title: Granzyme A binding to target cell proteins. Granzyme A binds to and cleaves nuc
A;Reference number: A40769; MUID:91317840
A;Accession: A40769
A;Molecule type: protein
A;Residues: 2-20,'X',22-24 <PAS>
R;Yang, T.H.; Tsai, W.H.; Lee, Y.M.; Lei, H.Y.; Lai, M.Y.; Chen, D.S.; Yeh, N.H.; Lee, S
Mol. Cell. Biol. 14, 6068-6074, 1994
A;Title: Purification and characterization of nucleolin and its identification as a tran
A;Reference number: A56240; MUID:94344117
A;Accession: A56240
A;Molecule type: protein
A;Residues: 2-19;558-567 <YAN>
R;Bourbon, H.
Gene 68, 73-84, 1988
A;Title: Sequence and structure of the nucleolin promoter in rodents: Characterization c
A;Reference number: I48118; MUID:89121496
A;Accession: I84688
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-44 <RES>
A;Cross-references: GB:M22089; NID:g200111; PIDN:AAA39841.1; PID:g554246
C;Comment: This housekeeping protein is involved in the synthesis, packaging, and matura
C;Genetics:
A;Introns: 6/3; 45/3; 211/1; 273/1; 302/1; 349/2; 391/1; 432/3; 484/1; 525/2; 566/1; 608
C;Superfamily: nucleolin; ribonucleoprotein repeat homology
C;Keywords: DNA binding; duplication; nucleus; phosphoprotein; RNA binding; transcriptio
F;310-375/Domain: ribonucleoprotein repeat homology <RRM1>
F;311-316/Region: RNA-binding RNP2 motif
F;349-356/Region: RNA-binding RNP1 motif
F;396-458/Domain: ribonucleoprotein repeat homology <RRM2>
F;397-402/Region: RNA-binding RNP2 motif
F;431-438/Region: RNA-binding RNP1 motif
F;488-551/Domain: ribonucleoprotein repeat homology <RRM3>
F;489-494/Region: RNA-binding RNP2 motif
F;524-531/Region: RNA-binding RNP1 motif
F;570-634/Domain: ribonucleoprotein repeat homology <RRM4>
F;571-576/Region: RNA-binding RNP2 motif
F;607-614/Region: RNA-binding RNP1 motif


  Query Match             13.0%;  Score 206.5;  DB 1;  Length 707;
  Best Local Similarity   33.0%;  Pred. No. 0.00043;
  Matches   65;  Conservative   15;  Mismatches   54;  Indels   63;  Gaps    9;

Qy      97 EQDDFTSSDDEDVPEAVPAPAPTAVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDE 156
           |:|:  | :|||  |      |            :||   || || |   || |  | ||
Db     151 EEDEDDSDEDEDDEEEDEFEPP------------IVKG-VKP-AKAAPAAPASEDEEDDE 196

Qy     157 -EDESDDEDESEEDDDSEKGMDV----------------------------DEDDSD-D 185
            ||: :|:|| ||||  |: |::                            |||| | |
Db     197 DEDDEEDDDEEEEDDSEEEVMEITTAKGKKTPAKVVPMKAKSVAEEEDDEEEDEDDEDED 256

Qy     186 DEEEDSED---EEEEETPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTEKKK 242
           ||||| ||   ||||| | |  |  :|:           :  |   || :  || |
Db     257 DEEEDDEDDDEEEEEEEPVKAAPGKRKK------------EMTKQKEAPEAKKQKVE--- 301

Qy     243 GGHTATPHPAKKGGKSP 259
           |    ||     |  :|
Db     302 GSEPTTPFNLFIGNLNP 318


RESULT  11
A27441
nucleolin - Chinese hamster
N;Alternate names: nonchromosomal protein C23; nucleolar protein C23
C;Species: Cricetulus griseus (Chinese hamster)
C;Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 23-Jul-1999
C;Accession: A27441; A24808; A24070
R;Lapeyre, B.; Bourbon, H.; Amalric, F.
Proc. Natl. Acad. Sci. U.S.A. 84, 1472-1476, 1987
A;Title: Nucleolin, the major nucleolar protein of growing eukaryotic cells: an unusu
A;Reference number: A27441; MUID:87175501
A;Accession: A27441
A;Molecule type: mRNA
A;Residues: 1-713 <LAP>
R;Lapeyre, B.; Amalric, F.; Ghaffari, S.H.; Venkatarama Rao, S.V.; Dumbar, T.S.; Olso
J. Biol. Chem. 261, 9167-9173, 1986
A;Title: Protein and cDNA sequence of a glycine-rich, dimethylarginine-containing reg
A;Reference number: A24808; MUID:86250857
A;Accession: A24808
A;Molecule type: mRNA
A;Residues: 522-541,'G',543-544,'R',546-713 <LA2>
A;Cross-references: GB:M13757; NID:g191004; PIDN:AAA36961.1; PID:g387047
R;Lapeyre, B.; Caizergues-Ferrer, M.; Bouche, G.; Amalric, F.
Nucleic Acids Res. 13, 5805-5816, 1985
A;Title: Cloning of cDNA encoding a 100 kDa nucleolar protein (nucleoline) of Chinese
A;Reference number: A24070; MUID:85297777
A;Accession: A24070
A;Molecule type: mRNA
A;Residues: 179-238 <LA3>
C;Superfamily: nucleolin; ribonucleoprotein repeat homology
C;Keywords: DNA binding; nucleus; phosphoprotein
F;308-373/Domain: ribonucleoprotein repeat homology <RRM1>
F;394-456/Domain: ribonucleoprotein repeat homology <RRM2>
F;486-549/Domain: ribonucleoprotein repeat homology <RRM3>
F;572-636/Domain: ribonucleoprotein repeat homology <RRM4>


  Query Match             13.0%;  Score 206;  DB 2;  Length 713;
  Best Local Similarity   23.2%;  Pred. No. 0.00046;
  Matches   66;  Conservative   31;  Mismatches   81;  Indels  106;  Gaps    8;

Qy      97 EQDDFTSSDDEDVPE----------------------AVPAPA---------------- 117
           |::| :| ::  :|:                      ||| ||
Db      34 EEEDDSSGEEVVIPQKKGKKATATPAKKVVVSQTKKVAVPTPAKKAAVTPGKKAAATPAK 93

Qy     118 -----PTAVTANGNAGAAVVKADTKPKAKPAEVKPAE-------EKPESDEEDESDDEDE 165
                 ||   |  ||   ||       |   | ||:        | |  :||| ||:||
Db      94 KAVTPAKAVATPGKKGATQAKALVATPGKKGAVTPAKGAKNGKNAKKEDSDEDEDDDDDE 153

Qy     166 SEEDDDSEKGMDVDEDD-------------------SDDDEEEDSEDEEEEETPKKPEP 205
            : |:| |   | :||:                  |:|||| |:|||:|     |
Db     154 DDSDEDEE---DEEEDEFEPPVVKGKQGKVAAAAPASEDEDEEEDEEEEEEDE-----EE 205

Qy     206 INKKRPNESVSKTPVSGKKAKPAAAPAST-------------------------PQKTE 239
           :     |::  ||  ||||     |                            :: |
Db     206 EDDSEEEEAMEITPAKGKKAPAKVVPVKAKNVAEEDDDEEEDEDEEEDEEEEEDEEEE 265

Qy     240 KKKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGGNNNKKPFN 283
           :::      | | |:  |      : |::  |   |: :  |||
Db     266 EEEEEEPVKPAPGKR-KKEMTKQKEVPEAKKQKVEGSESTTPFN 308


RESULT  12
A35804
nucleolin - human
N;Alternate names: phosphoprotein pp100; protein B50; protein C23
C;Species: Homo sapiens (man)
C;Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 23-Jul-1999
C;Accession: A35804; S04631; A48138; A55996
R;Srivastava, M.; McBride, O.W.; Fleming, P.J.; Pollard, H.B.; Burns, A.L.
J. Biol. Chem. 265, 14922-14931, 1990
A;Title: Genomic organization and chromosomal localization of the human nucleolin gen
A;Reference number: A35804; MUID:90368666
A;Accession: A35804
A;Molecule type: DNA
A;Residues: 1-707 <SRI>
A;Cross-references: GB:M60858; GB:J05584; NID:g189305; PIDN:AAA59954.1; PID:g189306
R;Srivastava, M.; Fleming, P.J.; Pollard, H.B.; Burns, A.L.
```

FEBS Lett. 250, 99-105, 1989
A;Title: Cloning and sequencing of the human nucleolin cDNA.
A;Reference number: S04631; MUID:89290043
A;Accession: S04631
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-707 <SR2>
R;Ishikawa, F.; Matunis, M.J.; Dreyfuss, G.; Cech, T.R.
Mol. Cell. Biol. 13, 4301-4310, 1993
A;Title: Nuclear proteins that bind the pre-mRNA 3' splice site sequence r(UUAG/G) and t
A;Reference number: A48138; MUID:93309464
A;Accession: A48138
A;Molecule type: protein
A;Residues: 458-474 <ISH>
A;Experimental source: HeLa cell nuclei
A;Note: sequence extracted from NCBI backbone (NCBIP:134645)
R;Jordan, P.; Heid, H.; Kinzel, V.; Kuebler, D.
Biochemistry 33, 14696-14706, 1994
A;Title: Major cell surface-located protein substrates of an ecto-protein kinase are hom
A;Reference number: A55996; MUID:95086063
A;Accession: A55996
A;Molecule type: protein
A;Residues: 231-236;349-362;399-403;458-461;655-656,'X',658-660 <JOR>
A;Experimental source: surface-labelled HeLa cells
C;Genetics:
A;Gene: GDB:NCL
A;Cross-references: GDB:125908; OMIM:164035
A;Map position: 2q12-2qter
C;Superfamily: nucleolin; ribonucleoprotein repeat homology
C;Keywords: DNA binding; nucleolus; nucleus; phosphoprotein; RNA binding
F;308-373/Domain: ribonucleoprotein repeat homology <RRM1>
F;394-456/Domain: ribonucleoprotein repeat homology <RRM2>
F;487-550/Domain: ribonucleoprotein repeat homology <RRM3>
F;573-634/Domain: ribonucleoprotein repeat homology <RRM4>

Query Match 12.9%; Score 205; DB 2; Length 707;
Best Local Similarity 27.2%; Pred. No. 0.00052;
Matches 74; Conservative 29; Mismatches 95; Indels 74; Gaps 9;

```
Qy      6 VAVTP-KNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGG------AKLVIGTLS 58
          | |:| |     || : : |   : :     |   :   :|| |       | :: |
Db     64 VVVSPTKKVAVATPAKKAAVTPGKKAAATPAKKTVTPAKAVTTPGKKGATPGKALVATPG 123

Qy     59 QDKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDF--TSSDDEDVPEAVPAP 116
          :             |   :   | |   :     |   | ||      ||||  |
Db    124 K------------KGAAIPAKGAKNGKNAKKEDSDEEEDDDSEEDEEDDEDEDEDEDEI 170

Qy    117 APTAVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEE-------- 168
           | |: |   | ||          ||     :|  | ||:|: |:||:|||
Db    171 EPAAMKA---AAAA----------PASEDEDDEDDEDDEDDDDDEEDDSEEEAMETTPA 216

Qy    169 ------------------DDDSEKGMDVDEDDSDDDEEEDSEDEEEEETPKKPEPINKK 209
                           :|: |   | |:|| :|||::| ||:|||| :: ||
Db    217 KGKKAAKVVPVKAKNVAEDEDEEEDDEDEDDDDDEDDEDDDDEDDEEEEEEEEEEP---- 272

Qy    210 RPNESVSKTPVSGKK--AKPAAAPASTPQKTE 239
              | : |   ||  ||  ||| :  || |
Db    273 -----VKEAPGKRKKEMAKQKAAPEAKKQKVE 299
```

RESULT 13
JC4295
heat-shock protein - slime mold (Dictyostelium discoideum)
C;Species: Dictyostelium discoideum
C;Date: 14-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 29-Oct-1999
C;Accession: JC4295
R;de Maria, A.C.; Gomes, S.L.; Juliania, M.H.; Mazzarella, R.; Klein, C.
Gene 163, 163-164, 1995
A;Title: Cloning of a cDNA encoding a novel heat-shock protein from Dictyostelium discoi
A;Reference number: JC4295; MUID:96001265
A;Accession: JC4295
A;Molecule type: mRNA
A;Residues: 1-281 <DEM>
A;Cross-references: GB:L39778; NID:g662358; PIDN:AAA99510.1; PID:g662359
C;Comment: This protein has roles in growth, development and thermo-tolerance.
C;Keywords: heat shock; stress-induced protein

Query Match 12.6%; Score 200.5; DB 2; Length 281;
Best Local Similarity 21.6%; Pred. No. 0.00036;
Matches 72; Conservative 49; Mismatches 123; Indels 89; Gaps 10;

```
Qy      1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQD 60
          |:|:|  || :    :  :|  : |:::| :    :    | |:        ||  : :|
Db      2 MQFFGTIVTKEEPVNLELDEGDIFHLTKAIIHPKSQGKGKVYLTA-------VISLMEED 54

Qy     61 KFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTA 120
          :                                :|:||    |:|: |       |
Db     55 E---------------------------------MEEDDV--DDEEESPREDIVEIPIG 78

Qy    121 VTANG----------NAGAAVVKADTKPKAKPAEVKPAE------EKPESDEEDESDDE 163
          :   |          | |  :|: : : :    |           |:   |:|| |:
Db     79 ILEAGKIDQIDLNLHYNFG-QIVRFELQAENGAGYVVALSGSVITMEQGGCDDEDCDDEH 137

Qy    164 DESEEDDDSEKGMDVDEDDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGK 223
            : |||:     ::| |:   | ::| || ::|| |:   |  ||     :|: | | |
Db    138 CINHEDDE-----EIDSDEEFGDSDQDEEDSDDEEIPQLIAPATKKGKITEISEVPESKK 192

Qy    224 KAKPAAAPASTPQKTE----------------KKKGGHTATPHPAKKGGKSPVNANQSP 266
          :  |       |:| :                :|    |   |||       || : |
Db    193 EKTPEPKKVPEPKKEQVKQPTQPQQKKAAAQQPEKANNKPAAASPAKPQNNQSKNAPKQP 252

Qy    267 KSGGQS-SGGNNNKKPFNSGKQFGGSNNKGSNK 298
          :   || :  ||||:| |  :     |||   |
Db    253 QQQQQSPAKNNNNKRPQNQNE-----NNKKKQK 280
```

RESULT 14
T09648
nucleolin homolog nuM1 - alfalfa
N;Alternate names: nuM1 protein
C;Species: Medicago sativa (alfalfa)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C;Accession: T09648
R;Boegre, L.; Jonak, C.; Mink, M.; Meskiene, I.; Traas, J.; Ha, D.T.C.; Swoboda, I.;
Plant Cell 8, 417-428, 1996
A;Title: Developmental and cell cycle regulation of Alfalfa nucMs1 a plant homolog of
A;Reference number: Z16796; MUID:96361876
A;Accession: T09648
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-635 <BOE>
A;Cross-references: EMBL:X88845; NID:g1279562; PIDN:CAA61298.1; PID:g1279563
C;Genetics:
A;Gene: nuM1
C;Superfamily: nucleolin; ribonucleoprotein repeat homology

Query Match 12.4%; Score 196.5; DB 2; Length 635;
Best Local Similarity 31.8%; Pred. No. 0.0013;
Matches 56; Conservative 24; Mismatches 61; Indels 35; Gaps 6;

```
Qy     92 KSPNIEQDDFTSSDDEDVPEAVPA----PAPTAVTANGNAGAAVVKA----DTKPKAKPA 143
          |  :   :|  |||:||   |  |      |||   |: :   :   ::     | || :|||
Db    211 KDDSSSDEDSESSDEEDKKPAAKASKNVSAPTKKAASSSDEESDEESDEDEDAKPVSKPA 270

Qy    144 EVKPAEEKPESDEEDESDD----------------EDESEEDDDSEKGMDVDEDDSDDD 186
            |    :|  || :|| ||                 | ||:  ||  | |::|:| || |
Db    271 AVAKKSKKDSSDSDDEDDDSSSDEDKKPVASKKEVSESESDSSDDDHK-MNIDKDSSDSD 329

Qy    187 EEEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTEKKK 242
          | |:  ::|  :||:|     | :  | |      :||  | |    :|| :|   |
Db    330 ESEEESEDEPLKTPQK-----KIKDVEMVD----AGKSGKKAPNTPATPNETSGSK 376
```

```
RESULT  15
T19431
hypothetical protein C25A1.10 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C;Accession: T19431
R;Mortimore, B.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19124
A;Accession: T19431
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-971 <WIL>
A;Cross-references: EMBL:Z81038; PIDN:CAB02755.1; GSPDB:GN00019; CESP:C25A1.10
A;Experimental source: clone C25A1
C;Genetics:
A;Gene: CESP:C25A1.10
A;Map position: 1
A;Introns: 38/3; 92/3; 201/3; 919/3
C;Superfamily: nucleolus-cytoplasm shuttle phosphoprotein


  Query Match             12.1%;  Score 192.5;  DB 2;  Length 971;
  Best Local Similarity   25.7%;  Pred. No. 0.0033;
  Matches   74;  Conservative   27;  Mismatches   90;  Indels   97;  Gaps   11;

Qy     103 SSDDEDVPEAVPAPA---PTAVTANGNAGAAVVKADTKPKAKP----AEVKPAEEKPESD 155
           ||||| | | ||||   | |     : :     : || |||    |  ||  :| ||
Db     346 SSDDEKKPVAKPAPAKATPKPVAKKAESSSDSSDDEKKPAAKPTPAKATPKPVAKKAES- 404

Qy     156 EEDESDDE--------------------DESEEDDDSE---------------KGMDV 178
             | ||||                    | | :  | |               |  :
Db     405 SSDSSDDEKKPVAKPTSAKATPKPAAKKADSSSDSSDDEAPAKKTPAKAAPKPASKKAES 464

Qy     179 DEDDSDDDE--------------------EEDSEDEEEEETPK------KPEPINKKR 210
             | |||::                    || |:|::   |       |:|: ||
Db     465 SSDSSDDEKPAAKSTPAKITPKPTAKKVASSSSDSSDDEKKPAAKPTPANATPKPVAKKA 524

Qy     211 PNESVSKTPVSGKKAKPAAAPAS---TPQKTEK--------------KKGGHTATPHPA 252
           : | |    |  : || | | |   ||:   |               ||    | | ||
Db     525 ESSSDS----SDDEKKPVAKPTSAKATPKPAAKKADLSSDFSDDEAPAKKTPAKAAPKPA 580

Qy     253 KKGGKSPVNANQSPKSGGQSSGGNNNKKPFNSGKQFGGSNNKGSNKGK 300
           |  :|  :::   |   :|:      ||  : |:   |::  |:  |
Db     581 SKKAESSSDSSDDEKPAAKSTPAKTTPKP--TAKKAASSSSDSSDDEK 626


Search completed: May  2, 2002, 22:07:31
Job time: 6679 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: May 2, 2002, 22:31:08 ; Search time 36.81 Seconds
(without alignments)
320.822 Million cell updates/sec

Title: US-09-645-337-8
Perfect score: 1589
Sequence: 1 MEFWGVAVTPKNATKVTPEE.....KQFGGSNNKGSNKGKGKGRA 305

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	311	19.6	412	1	FKB4_SPOFR	Q26486 spodoptera
2	240	15.1	357	1	FKB4_DROME	P54397 drosophila
3	215.5	13.6	694	1	NUCL_CHICK	P15771 gallus gall
4	211.5	13.3	712	1	NUCL_RAT	P13383 rattus norv
5	208	13.1	411	1	MP62_LYTPI	P91753 lytechinus
6	206.5	13.0	706	1	NUCL_MOUSE	P09405 mus musculu
7	206	13.0	713	1	NUCL_MESAU	P08199 mesocricetu
8	205	12.9	706	1	NUCL_HUMAN	P19338 homo sapien
9	200.5	12.6	281	1	HS32_DICDI	P54658 dictyosteli
10	188	11.8	704	1	NP14_RAT	P41777 rattus norv
11	188	11.8	723	1	SSRP_DROME	Q05344 drosophila
12	183.5	11.5	294	1	NPM_CHICK	P16039 gallus gall
13	178.5	11.2	517	1	T2FA_HUMAN	P35269 homo sapien
14	177	11.1	644	1	NFM_RABIT	P54938 oryctolagus
15	177	11.1	650	1	NUCL_XENLA	P20397 xenopus lae
16	175	11.0	524	1	T2FA_XENLA	Q04870 xenopus lae
17	173.5	10.9	361	1	FKB4_SCHPO	O74191 schizosacch
18	172	10.8	292	1	NPM_MOUSE	Q61937 mus musculu
19	172	10.8	678	1	GARP_PLAFF	P13816 plasmodium
20	170	10.7	1089	1	NMD2_YEAST	P38798 saccharomyc
21	169.5	10.7	543	1	TLP1_MOUSE	Q9z273 mus musculu
22	169	10.6	292	1	NPM_RAT	P13084 rattus norv
23	168.5	10.6	299	1	NPM_XENLA	P07222 xenopus lae
24	168.5	10.6	699	1	NP14_HUMAN	Q14978 homo sapien
25	168	10.6	411	1	FKB3_YEAST	P38911 saccharomyc
26	168	10.6	1002	1	YEMA_DROME	P25992 drosophila
27	167.5	10.5	915	1	NFM_HUMAN	P07197 homo sapien
28	167	10.5	1275	1	TRP_DROME	P19334 drosophila
29	166	10.4	304	1	CEC1_CAEEL	P34618 caenorhabdi
30	166	10.4	542	1	TUL1_HUMAN	O00294 homo sapien
31	164.5	10.4	472	1	YWIE_CAEEL	Q23525 caenorhabdi
32	164.5	10.4	589	1	RGP1_MOUSE	P46061 mus musculu
33	162.5	10.2	1109	1	TCF8_RAT	Q62947 rattus norv
34	162	10.2	940	1	MAZ4_SCHCO	P37938 schizophyll
35	161.5	10.2	217	1	YKR4_EBV	P30117 epstein-bar
36	161.5	10.2	562	1	SIS2_YEAST	P36024 saccharomyc
37	161.5	10.2	1020	1	NFH_HUMAN	P12036 homo sapien
38	161	10.1	233	1	SERA_ENTHI	P21138 entamoeba h
39	160.5	10.1	294	1	NPM_HUMAN	P06748 homo sapien
40	160.5	10.1	1217	1	AF4_MOUSE	O88573 mus musculu
41	160	10.1	845	1	NFM_RAT	P12839 rattus norv
42	160	10.1	1637	1	MRSP_STAAU	P80544 staphylococ
43	160	10.1	2175	1	HMCU_DROME	P10180 drosophila
44	158.5	10.0	716	1	PEP_DROME	P41073 drosophila
45	158	9.9	848	1	NFM_MOUSE	P08553 mus musculu

ALIGNMENTS

RESULT 1
FKB4_SPOFR
ID FKB4_SPOFR STANDARD; PRT; 412 AA.
AC Q26486;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE 46 kDa FK506-binding nuclear protein (Peptidyl-prolyl cis-trans
DE isomerase) (PPIase) (EC 5.2.1.8).
OS Spodoptera frugiperda (Fall armyworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Noctuoidea; Noctuidae; Amphipyrinae; Spodoptera.
OX NCBI_TaxID=7108;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95074110; PubMed=7527037;
RA Alnemri E.S., Fernandes-Alnemri T., Pomerenke K., Robertson N.M.,
RA Dudley K., Dubois G.C., Litwack G.;
RT "FKBP46, a novel Sf9 insect cell nuclear immunophilin that forms a
RT protein-kinase complex.";
RL J. Biol. Chem. 269:30828-30834(1994).
CC -!- FUNCTION: BINDS TO, AND IS INHIBITED BY FK506 AND RAPAMYCIN. BINDS
CC DOUBLE-STRANDED DNA IN VITRO. PPIASES ACCELERATE PROTEIN FOLDING.
CC -!- CATALYTIC ACTIVITY: CIS-TRANS ISOMERIZATION OF PROLINE IMIDIC
CC PEPTIDE BONDS IN OLIGOPEPTIDES.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- PTM: PHOSPHORYLATED BY A NUCLEAR KINASE IN THE PRESENCE OF MG2+
CC AND ATP.
CC -!- SIMILARITY: BELONGS TO THE FKBP-TYPE PPIASE FAMILY.
CC ---

CC ---
DR EMBL; U15038; AAA58962.1; -.
DR HSSP; P20071; 1FKJ.
DR InterPro; IPR001179; FKBP_PPIase.
DR Pfam; PF00254; FKBP; 1.
DR PROSITE; PS00453; FKBP_PPIASE_1; FALSE_NEG.
DR PROSITE; PS00454; FKBP_PPIASE_2; 1.
DR PROSITE; PS50059; FKBP_PPIASE_3; 1.
KW Isomerase; Rotamase; Nuclear protein; DNA-binding; Phosphorylation.
FT DOMAIN 91 112 ASP/GLU-RICH (HIGHLY ACIDIC).
FT DOMAIN 120 145 LYS-RICH (BASIC).
FT DOMAIN 152 216 ASP/GLU-RICH (HIGHLY ACIDIC).
FT DOMAIN 219 302 LYS-RICH (BASIC).
FT DOMAIN 324 412 PPIASE, FKBP-TYPE.
SQ SEQUENCE 412 AA; 45810 MW; F2A69159AEF4FE22 CRC64;

Query Match 19.6%; Score 311; DB 1; Length 412;

```
  Best Local Similarity   31.0%;  Pred. No. 2.8e-10;
  Matches  100;  Conservative   39;  Mismatches  116;  Indels   68;  Gaps   14;

Qy       3 FWGVAVTP-KNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQDK 61
           |||: : | |  |:|    :   |||||::| :    :   : | | |   :: || : |
Db       2 FWGLIMEPNKRYTQVV---EKPFHISQAAMDISTGDNDPCQVMVVVDGKNFLVCTLQKGK 58

Qy      62 FPQISFDLVFDKEFELSH-SGTKANVHFIGYKSPNIEQD--DFTSSDDEDVPEAVPAPAP 118
             |:  || |     :|  :  | |||  ||  |  |:|  |   :::|:  | |   |
Db      59 IIQVPLDLYFKSGDSVSFLTNGKCNVHLTGYLDPEFEEDLEDEEEAEEEEEEEAPPLVP 118

Qy     119 T---AVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEED----------ESDDED 164
                | |   |  ||    ||       ||   : |:||           ::|: |
Db     119 AKNKRKLENANDATANKKAKPDKKAGKNSAPAAESDSDDDDEDQLQKFLDGEDIDTDEND 178

Qy     165 ESEEDDDSEKGMDVDEDDSDDDEE----EDSEDEEEEETPKKPEPINKKRP--------- 211
           || : : | :| | ||:| |:|||    :| :|||||| |||     ||:|
Db     179 ESFKMNTSAEGDDSDEEDDDEDEEDEEDDDEDDEEEEEAPKK----KKKQPAAEQDSTLD 234

Qy     212 --NESV--SKTPVSGK----------KAKPAAAPASTP------QKTEKKKGGHTATPH 250
             |||  ||   | |           | :|       |      || ||||
Db     235 TSKESVDMSKLSKSQKRRLKKKLQQQAKQQPQVNGVDKPKKEEPQQKAEKKK-------- 286

Qy     251 PAKKGGKSPVNANQSPK-SGGQS 272
           |  |  ::||   :  : :|| |
Db     287 PEAKKEEAPVEKKEKKQIAGGVS 309


RESULT  2
FKB4_DROME
ID   FKB4_DROME     STANDARD;      PRT;   357 AA.
AC   P54397; Q9VF88;
DT   01-OCT-1996 (Rel. 34, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   39 kDa FK506-binding nuclear protein (Peptidyl-prolyl cis-trans
DE   isomerase) (PPiase) (EC 5.2.1.8).
GN   FK506-BP1 OR FKBP39 or CG6226.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CANTON-S;
RX   MEDLINE=95278752; PubMed=7538962;
RA   Theopold U., Dal Zotto L., Hultmark D.;
RT   "FKBP39, a Drosophila member of a family of proteins that bind the
RT   immunosuppressive drug FK506.";
RL   Gene 156:247-251(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- FUNCTION: PPIASES ACCELERATE THE FOLDING OF PROTEINS. FKBP506
CC       MAY FUNCTION IN A SIGNAL TRANSDUCTION CASCADE DURING EARLY
CC       DEVELOPMENT.
CC   -!- CATALYTIC ACTIVITY: CIS-TRANS ISOMERIZATION OF PROLINE IMIDIC
CC       PEPTIDE BONDS IN OLIGOPEPTIDES.
CC   -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC   -!- TISSUE SPECIFICITY: UBIQUITOUSLY EXPRESSED, HIGHEST LEVELS IN
CC       OVARY.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED DURING ALL STAGES OF DEVELOPMENT
CC       WITH HIGHEST EXPRESSION IN EARLY EMBRYO.
CC   -!- SIMILARITY: BELONGS TO THE FKBP-TYPE PPIASE FAMILY.
CC   -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-3 IS THE INITIATOR.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z46894; CAA86996.1; -.
DR   EMBL; AE003708; AAF55172.1; ALT_SEQ.
DR   HSSP; P27124; 1ROT.
DR   FlyBase; FBgn0013269; FK506-bp1.
DR   InterPro; IPR001179; FKBP_PPIase.
DR   Pfam; PF00254; FKBP; 1.
DR   PROSITE; PS00453; FKBP_PPIASE_1; FALSE_NEG.
DR   PROSITE; PS00454; FKBP_PPIASE_2; 1.
DR   PROSITE; PS50059; FKBP_PPIASE_3; 1.
KW   Isomerase; Rotamase; Nuclear protein.
FT   DOMAIN       89     99       ASP/GLU-RICH (HIGHLY ACIDIC).
FT   DOMAIN      119    183       ASP/GLU-RICH (HIGHLY ACIDIC).
FT   DOMAIN      186    247       LYS-RICH (BASIC).
FT   DOMAIN      269    357       PPIASE, FKBP-TYPE.
FT   CONFLICT    187    187       A -> R (IN REF. 1).
SQ   SEQUENCE   357 AA;  39343 MW;  EF0AB7831738BB30 CRC64;


  Query Match             15.1%;  Score 240;  DB 1;  Length 357;
  Best Local Similarity   26.9%;  Pred. No. 1.4e-06;
  Matches   95;  Conservative   48;  Mismatches  110;  Indels  100;  Gaps   18;

Qy       3 FWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQDKF 62
           |||: : |:   | :       |||  :||     |:   | :     : :: |:::
Db       4 FWGLNMKPER--KYSQTIIKSPHISGVALD----KGQEAKLYLAAEKQEYIVATVTK-AI 56

Qy      63 PQISFDLVFDK-EFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTAV 121
           ||:: || | | :  : ::   |:|  :||        |  | ||||  :         :
Db      57 PQVALDLNFSKGDRIMFYTAGDASVSLLGY------LHDIDSEDDEDDDQ---------M 101

Qy     122 TANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDES-------------E 167
           |      : :|   |        :  |:: || |||| | :|:|             |
```

```
Db    102 TIENLLNSKAIKNSKKS-------EDDEDENESGEEDEEDTDDDSQIIEEYESFLENGEE 154

Qy    168 EDDDSEKGMDVDED--DSDDDEEEDSEDEE-EEETPKKPEPINKKRPNESVSKTPVSGKK 224
          ||||     |||||  :| :::|:||:| | ||| ||    : |  |  |  |   |||:
Db    155 EDDD-----DVDEDNEESGEEDEQDSDDSEAEEEQPK--AKVAKLSPGASAKK---SGKE 204

Qy    225 AKPAAAPASTPQKTEKKKGGHTATPHPAKKGGKSPVNANQSPKS---------------- 268
              |      |: :||:      |   |: |:   | | | |
Db    205 QNGVAKKEEAKQQQKKKE-----KPEAKKEQPKAKEPAKQQPASKDPRTITGGVKIVDQV 259

Qy    269 ---GGQSSGG----------NNNKKPFNS---GKQF-----GGSNNKGSNKG 299
            | ::  |           :| | |:|   || |     ||   || : |
Db    260 VGKGEEAKQGKRVSVYYIGRLQSNNKTFDSLLKGKPFKPALGGGEVIKGWDVG 312


RESULT   3
NUCL_CHICK
ID   NUCL_CHICK     STANDARD;      PRT;   694 AA.
AC   P15771;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Nucleolin (Protein C23).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90206792; PubMed=2320420;
RA   Maridor G., Nigg E.A.;
RT   "cDNA sequences of chicken nucleolin/C23 and NO38/B23, two major
RT   nucleolar proteins.";
RL   Nucleic Acids Res. 18:1286-1286(1990).
RN   [2]
RP   DISCUSSION OF SEQUENCE.
RX   MEDLINE=90304215; PubMed=2114180;
RA   Maridor G., Krek W., Nigg E.A.;
RT   "Structure and developmental expression of chicken nucleolin and
RT   NO38: coordinate expression of two abundant non-ribosomal nucleolar
RT   proteins.";
RL   Biochim. Biophys. Acta 1049:126-133(1990).
RN   [3]
RP   SEQUENCE OF 407-694 FROM N.A.
RX   MEDLINE=89119560; PubMed=2914325;
RA   Borer R.A., Lehner C.F., Eppenberger H.M., Nigg E.A.;
RT   "Major nucleolar proteins shuttle between nucleus and cytoplasm.";
RL   Cell 56:379-390(1989).
CC   -!- FUNCTION: NUCLEOLIN IS THE MAJOR NUCLEOLAR PROTEIN OF GROWING
CC       EUKARYOTIC CELLS. IT IS FOUND ASSOCIATED WITH INTRANUCLEOLAR
CC       CHROMATIN AND PRERIBOSOMAL PARTICLES. IT INDUCES CHROMATIN
CC       DECONDENSATION BY BINDING TO HISTONE H1. IT IS THOUGHT TO PLAY A
CC       ROLE IN PRE-RRNA TRANSCRIPTION AND RIBOSOME ASSEMBLY.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
CC   -!- SIMILARITY: CONTAINS 4 RNA RECOGNITION MOTIFS (RRM).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X17199; CAA35060.1; -.
DR   EMBL; M21791; AAA48983.1; -.
DR   PIR; S08414; DNCHNL.
DR   HSSP; P09651; 1UP1.
DR   InterPro; IPR000504; RRM.
DR   Pfam; PF00076; rrm; 4.
DR   SMART; SM00360; RRM; 4.
DR   PROSITE; PS50102; RRM; 4.
DR   PROSITE; PS00030; RRM_RNP_1; 3.
KW   Nuclear protein; Phosphorylation; Methylation; DNA-binding; Repeat;
KW   RNA-binding.
FT   DOMAIN      124    141       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      170    192       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      217    247       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      281    357       RNA-BINDING (RRM) 1.
FT   DOMAIN      371    445       RNA-BINDING (RRM) 2.
FT   DOMAIN      461    535       RNA-BINDING (RRM) 3.
FT   DOMAIN      553    628       RNA-BINDING (RRM) 4.
FT   DOMAIN      630    682       ARG/GLY/PHE-RICH.
FT   DOMAIN       55     90       5 X 7 AA TANDEM REPEATS OF X-T-P-X-K-K-X.
FT   REPEAT       55     61       1.
FT   REPEAT       62     68       2.
FT   REPEAT       69     75       3.
FT   REPEAT       76     82       4.
FT   REPEAT       84     90       5.
FT   MOD_RES     116    116       PHOSPHORYLATION (BY SIMILARITY).
FT   MOD_RES     136    136       PHOSPHORYLATION (BY SIMILARITY).
FT   MOD_RES     171    171       PHOSPHORYLATION (BY SIMILARITY).
FT   CONFLICT    419    419       A -> R (IN REF. 3).
FT   CONFLICT    520    520       N -> T (IN REF. 3).
SQ   SEQUENCE   694 AA;  75640 MW;  7996C504BE9459A1 CRC64;


  Query Match             13.6%;  Score 215.5;  DB 1;  Length 694;
  Best Local Similarity   28.1%;  Pred. No. 5.2e-05;
  Matches   85;  Conservative   32;  Mismatches  102;  Indels   83;  Gaps   12;

Qy      6 VAVTPKNATK----VTPEEDSLVHISQASLDCTVKSGESVVL--------SVT------- 46
          :| |||| |      | :       : | |    ||| |::         :||
Db      4 LAKTPKNQMKQKKMAPPPKKVEESEEEESSDLEESSGEEVMVPPKKQQKAAVTPAKKAAT 63

Qy     47 ------------VGGAKLVIGTLSQDKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSP 94
                      |  ||  : |       |          |:  :   | |   :    :|
Db     64 PAKKAATPAKKAVTPAKKAVAT------PAKKAVAPSPKKAAVVGKGAKNGKNAKKEESE 117

Qy     95 NIEQDDFTSSDDEDVPEAVPAPAPTAVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPES 154
            ::||    :|||  |         ||       ||    |   |  ||| | ||  | ||
Db    118 EEDEDDEDDEEDEDEEEESDEEEEPAVPVK---PAAKKSAAAVPAKKPA-VVPA--KQES 171

Qy    155 DEEDESDDEDESEEDDDSE-KGMDV-------------------------DEDDSDDDE 187
          :||:| |||:| ||||:|| : ||                          ||:| |:||
Db    172 EEEEEEDDEEEDEEDDESEDEAMDTTPAPVKKPTPAKATPAKAKAESEDEEDEEDEDDE 231

Qy    188 EEDSEDEEEEETPKKPEPINK---KRPNESVSKTPVSGKKAKPAAAPAS---------TP 235
          |:: ::||:||  :  :|: :   ||  |  :|:    || |    |||          ||
Db    232 EDEDDEEEDEEESEDEKPVKEAPGKRKKEMANKSAPEAKK-KKTETPASAFSLFVKNLTP 290

Qy    236 QK 237
          |
Db    291 TK 292


RESULT   4
NUCL_RAT
ID   NUCL_RAT       STANDARD;      PRT;   712 AA.
AC   P13383;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Nucleolin (Protein C23).
GN   NCL OR NUC.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90269607; PubMed=2347493;
RA   Bourbon H.-M., Amalric F.;
RT   "Nucleolin gene organization in rodents: highly conserved sequences
RT   within three of the 13 introns.";
```

```
RL   Gene 88:187-196(1990).
RN   [2]
RP   SEQUENCE OF 1-44 FROM N.A.
RX   MEDLINE=89121496; PubMed=2906027;
RA   Bourbon H.-M., Prudhomme M., Amalric F.;
RT   "Sequence and structure of the nucleolin promoter in rodents:
RT   characterization of a strikingly conserved CpG island.";
RL   Gene 68:73-84(1988).
CC   -!- FUNCTION: NUCLEOLIN IS THE MAJOR NUCLEOLAR PROTEIN OF GROWING
CC       EUKARYOTIC CELLS. IT IS FOUND ASSOCIATED WITH INTRANUCLEOLAR
CC       CHROMATIN AND PRERIBOSOMAL PARTICLES. IT INDUCES CHROMATIN
CC       DECONDENSATION BY BINDING TO HISTONE H1. IT IS THOUGHT TO PLAY A
CC       ROLE IN PRE-RRNA TRANSCRIPTION AND RIBOSOME ASSEMBLY.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
CC   -!- SIMILARITY: CONTAINS 4 RNA RECOGNITION MOTIFS (RRM).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M55022; AAA41732.1; -.
DR   EMBL; M55015; AAA41732.1; JOINED.
DR   EMBL; M55017; AAA41732.1; JOINED.
DR   EMBL; M55020; AAA41732.1; JOINED.
DR   EMBL; M22090; AAA41733.1; -.
DR   PIR; JH0148; JH0148.
DR   HSSP; P09651; 1HA1.
DR   InterPro; IPR000504; RRM.
DR   Pfam; PF00076; rrm; 4.
DR   SMART; SM00360; RRM; 4.
DR   PROSITE; PS50102; RRM; 4.
DR   PROSITE; PS00030; RRM_RNP_1; 3.
KW   Nuclear protein; Phosphorylation; Methylation; DNA-binding; Repeat;
KW   RNA-binding.
FT   INIT_MET      0      0
FT   DOMAIN      142    167       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      187    215       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      241    274       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      310    386       RNA-BINDING (RRM) 1.
FT   DOMAIN      396    469       RNA-BINDING (RRM) 2.
FT   DOMAIN      488    562       RNA-BINDING (RRM) 3.
FT   DOMAIN      574    649       RNA-BINDING (RRM) 4.
FT   DOMAIN      651    702       ARG/GLY/PHE-RICH.
FT   DOMAIN       57    134       8 X 8 AA TANDEM REPEATS OF X-T-P-X-K-K-
FT                                X-X.
FT   REPEAT       57     64       1.
FT   REPEAT       74     81       2.
FT   REPEAT       82     89       3.
FT   REPEAT       90     97       4.
FT   REPEAT       98    103       5 (INCOMPLETE).
FT   REPEAT      104    111       6.
FT   REPEAT      119    126       7.
FT   REPEAT      127    134       8.
SQ   SEQUENCE   712 AA;  77016 MW;  68774A214E550F90 CRC64;


  Query Match           13.3%;  Score 211.5;  DB 1;  Length 712;
  Best Local Similarity  29.4%;  Pred. No. 8.7e-05;
  Matches   57;  Conservative   27;  Mismatches   55;  Indels   55;  Gaps    7;

Qy      97 EQDDFTSSDDEDVPEAVPAPAPTAVTANGNAGAAVVKADTKPKAKPAEVKPA-------E 149
           |:|:  | :|||  :    |              |||       |||:  ||       :
Db     150 EEDEDDSDEDEDEEDEFEPP-------------VVKG-----VKPAKAAPAAPASEDED 190

Qy     150 EKPESDEEDESDDEDESEEDDDSEKGMDV----------------------DEDDSDDD 186
           |: : ||:|: |||:| ||||  |: |::                     :||| ||:
Db     191 EEDDDDEDDDDDDEEEEEEDDSEEEVMEITPAKGKKTPAKVVPVKAKSVAEEEEDDEDDE 250

Qy     187 EEEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQ-KTEKKKGGH 245
```

```
           :||: ||||:||     |  :::   | |   |  ||: |        |: | :| :|
Db     251 DEEEDEDEEDEE---DDEDEDEEEEEEPVKAAP--GKRKKEMTKQKEAPEAKKQKIEGSE 305

Qy     246 TATPHPAKKGGKSP 259
             ||     |  :|
Db     306 PTTPFNLFIGNLNP 319


RESULT   5
MP62_LYTPI
ID   MP62_LYTPI     STANDARD;      PRT;   411 AA.
AC   P91753;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Mitotic apparatus protein P62.
OS   Lytechinus pictus (Painted sea urchin).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Echinacea; Temnopleuroida; Toxopneustidae;
OC   Lytechinus.
OX   NCBI_TaxID=7653;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97166213; PubMed=9013612;
RA   Ye X., Sloboda R.D.;
RT   "Molecular characterization of p62, a mitotic apparatus protein
RT   required for mitotic progression.";
RL   J. Biol. Chem. 272:3606-3614(1997).
CC   -!- FUNCTION: REQUIRED FOR MITOTIC PROGRESSION. BINDS TO CHROMATIN.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- PTM: PHOSPHORYLATED BY CAM-KINASE II IN VITRO.
CC   -!- SIMILARITY: BELONGS TO THE NUCLEOPLASMIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U76750; AAB47481.1; -.
DR   InterPro; IPR004301; Nucleoplasmin.
DR   Pfam; PF03066; Nucleoplasmin; 1.
KW   Nuclear protein; Phosphorylation; Mitosis.
FT   DOMAIN      147    156       POLY-GLU.
FT   DOMAIN      192    200       POLY-GLU.
FT   DOMAIN      262    267       POLY-GLU.
FT   DOMAIN      272    284       POLY-GLU.
FT   DOMAIN      293    300       POLY-ASP.
FT   DOMAIN      313    319       POLY-GLU.
SQ   SEQUENCE   411 AA;  46385 MW;  D2DE04293C362254 CRC64;


  Query Match           13.1%;  Score 208;  DB 1;  Length 411;
  Best Local Similarity  25.5%;  Pred. No. 8e-05;
  Matches   93;  Conservative   36;  Mismatches   97;  Indels  138;  Gaps   17;

Qy       3 FWGVAVT-PKNATKVTP----------EEDSLVH---ISQASLDCTVK---------- 36
           |||  ::  |   |  |           ||||: |   : || |    |
Db       6 FWGATLSKDKKIFKWDPESDFLDDEDDDEEDSISHFLFLKQAVLGVNAKDDDRNVIEVET 65

Qy      37 ---SGESVV---LSVTVG--------------GAKLVIGTLSQDKFPQISFDLVFDKEF 75
              ||:|:   ||: :|                || :|:        | : ||  |:||
Db      66 INFDGETVIQPLLSLRLGLNESTNLDIGLQPPVTFKLALGSGPVYLSGQHALDLQEDEEF 125

Qy      76 ELSHSGTKANVHFIG----------------------------------------YKSPN 95
              | :|  : :|                                          |
Db     126 GKDFEGAEA--YEVGDEDLEDEDEGEEDEEEEETPKKGSPKRIVKKIAAVKGRMKGKGDE 183

Qy      96 IEQDDFTSSDDEDVPEAVPA------PAPTAVTANGNAG--AAVVKADTKPKAKP---AE 144
           :::|:    ::|:  | :        |||   :| | |   | | |  |  :  |    |
Db     184 LDEDEDDDEEEEEEEEEIQTAKGKKRPAP---SAKGPAKKLAKVDKDGTSKRKVPNGSVE 240
```

```
Qy     145 VKPAEEKPESDEED------------------ESDDEDESEEDDDSEKGMDVDEDDSDD 185
           |  :  | ||||                    | :||:| ||:|| |  :  |:|| ||
Db     241 NGHAIDDDEDDEEDEDYKVGDEEEEEEATSGEEEEEDEEEEEEEEDDEEMALG-DDDDEDD 299

Qy     186 DEEE-------DSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQKT 238
           |||:       | |||||||     :|  |       :|  |:|  |||   | |   |
Db     300 DEEDDEDEEGMDDEDEEEEEDSSPVKPAKK-------AKGKVNG-TAKPKGTPKSQANKG 351

Qy     239 EKKK 242
             |:|
Db     352 MKEK 355


RESULT   6
NUCL_MOUSE
ID   NUCL_MOUSE     STANDARD;      PRT;   706 AA.
AC   P09405; Q61991;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Nucleolin (Protein C23).
GN   NCL OR NUC.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C;
RX   MEDLINE=88316930; PubMed=3137346;
RA   Bourbon H., Lapeyre B., Amalric F.;
RT   "Structure of the mouse nucleolin gene. The complete sequence reveals
RT   that each RNA binding domain is encoded by two independent exons.";
RL   J. Mol. Biol. 200:627-638(1988).
RN   [2]
RP   SEQUENCE OF 1-44 FROM N.A.
RX   MEDLINE=89121496; PubMed=2906027;
RA   Bourbon H.M., Prudhomme M., Amalric F.;
RT   "Sequence and structure of the nucleolin promoter in rodents:
RT   characterization of a strikingly conserved CpG island.";
RL   Gene 68:73-84(1988).
RN   [3]
RP   SEQUENCE OF 1-23.
RX   MEDLINE=91317840; PubMed=1860869;
RA   Pasternack M.S., Bleier K.J., McInerney T.N.;
RT   "Granzyme A binding to target cell proteins. Granzyme A binds to and
RT   cleaves nucleolin in vitro.";
RL   J. Biol. Chem. 266:14703-14708(1991).
CC   -!- FUNCTION: NUCLEOLIN IS THE MAJOR NUCLEOLAR PROTEIN OF GROWING
CC       EUKARYOTIC CELLS. IT IS FOUND ASSOCIATED WITH INTRANUCLEOLAR
CC       CHROMATIN AND PRERIBOSOMAL PARTICLES. IT INDUCES CHROMATIN
CC       DECONDENSATION BY BINDING TO HISTONE H1. IT IS THOUGHT TO PLAY A
CC       ROLE IN PRE-RRNA TRANSCRIPTION AND RIBOSOME ASSEMBLY.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
CC   -!- SIMILARITY: CONTAINS 4 RNA RECOGNITION MOTIFS (RRM).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X07699; CAA30538.1; -.
DR   EMBL; M22089; AAA39841.1; -.
DR   PIR; A29958; DNMS.
DR   HSSP; P09651; 1HA1.
DR   MGD; MGI:97286; Ncl.
DR   InterPro; IPR000504; RRM.
DR   Pfam; PF00076; rrm; 4.
DR   SMART; SM00360; RRM; 4.
DR   PROSITE; PS50102; RRM; 4.
DR   PROSITE; PS00030; RRM_RNP_1; 3.
KW   Nuclear protein; Phosphorylation; Methylation; DNA-binding; Repeat;
KW   RNA-binding.
FT   INIT_MET      0      0
FT   DOMAIN      142    169       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      189    214       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      240    272       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      308    384       RNA-BINDING (RRM) 1.
FT   DOMAIN      394    467       RNA-BINDING (RRM) 2.
FT   DOMAIN      486    560       RNA-BINDING (RRM) 3.
FT   DOMAIN      568    643       RNA-BINDING (RRM) 4.
FT   DOMAIN      645    696       ARG/GLY/PHE-RICH.
FT   MOD_RES     144    144       PHOSPHORYLATION (BY SIMILARITY).
FT   MOD_RES     156    156       PHOSPHORYLATION (BY SIMILARITY).
FT   MOD_RES     188    188       PHOSPHORYLATION (BY SIMILARITY).
FT   DOMAIN       57    134       8 X 8 AA TANDEM REPEATS OF X-T-P-X-K-K-
FT                               X-X.
FT   REPEAT       57     64       1.
FT   REPEAT       74     81       2.
FT   REPEAT       82     89       3.
FT   REPEAT       90     97       4.
FT   REPEAT       98    103       5 (INCOMPLETE).
FT   REPEAT      104    111       6.
FT   REPEAT      119    126       7.
FT   REPEAT      127    134       8.
SQ   SEQUENCE   706 AA;  76592 MW;  89505EE39C89F832 CRC64;


  Query Match             13.0%;  Score 206.5;  DB 1;  Length 706;
  Best Local Similarity   33.0%;  Pred. No. 0.00016;
  Matches   65;  Conservative   15;  Mismatches   54;  Indels   63;  Gaps    9;

Qy      97 EQDDFTSSDDEDVPEAVPAPAPTAVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDE 156
           |:|:  | :|||  |      |            :||   || || |   || |  | ||
Db     150 EEDEDDSDEDEDDEEEDEFEPP-----------IVKG-VKP-AKAAPAAPASEDEEDDE 195

Qy     157 -EDESDDEDESEEDDDSEKGMDV----------------------------DEDDSD-D 185
            ||: :|:|| ||||  |: |::                            |||| | |
Db     196 DEDDEEDDDEEEEDDSEEEVMEITTAKGKKTPAKVVPMKAKSVAEEEDDEEEDEDDEDED 255

Qy     186 DEEEDSED---EEEEETPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTEKKK 242
           ||||| ||   ||||| | |  |  :|:          :  |   || :  || |
Db     256 DEEEDDEDDDEEEEEEEPVKAAPGKRKK-----------EMTKQKEAPEAKKQKVE--- 300

Qy     243 GGHTATPHPAKKGGKSP 259
           |    ||     |  :|
Db     301 GSEPTTPFNLFIGNLNP 317


RESULT   7
NUCL_MESAU
ID   NUCL_MESAU     STANDARD;      PRT;   713 AA.
AC   P08199;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Nucleolin (Protein C23).
GN   NCL.
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE OF 35-713 FROM N.A., AND SEQUENCE OF 1-34.
RX   MEDLINE=87175501; PubMed=3470736;
RA   Lapeyre B., Bourbon H., Amalric F.;
RT   "Nucleolin, the major nucleolar protein of growing eukaryotic cells:
RT   an unusual protein structure revealed by the nucleotide sequence.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:1472-1476(1987).
RN   [2]
RP   FUNCTION.
```

```
RX   MEDLINE=88312631; PubMed=3409881;
RA   Erard M.S., Belenguer P., Caizergues-Ferrer M., Pantaloni A.,
RA   Amalric F.;
RT   "A major nucleolar protein, nucleolin, induces chromatin
RT   decondensation by binding to histone H1.";
RL   Eur. J. Biochem. 175:525-530(1988).
CC   -!- FUNCTION: NUCLEOLIN IS THE MAJOR NUCLEOLAR PROTEIN OF GROWING
CC       EUKARYOTIC CELLS. IT IS FOUND ASSOCIATED WITH INTRANUCLEOLAR
CC       CHROMATIN AND PRERIBOSOMAL PARTICLES. IT INDUCES CHROMATIN
CC       DECONDENSATION BY BINDING TO HISTONE H1. IT IS THOUGHT TO PLAY A
CC       ROLE IN PRE-RRNA TRANSCRIPTION AND RIBOSOME ASSEMBLY.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
CC   -!- SIMILARITY: CONTAINS 4 RNA RECOGNITION MOTIFS (RRM).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M15825; AAA36966.1; -.
DR   PIR; A27441; A27441.
DR   HSSP; P09651; 1HA1.
DR   InterPro; IPR000504; RRM.
DR   Pfam; PF00076; rrm; 4.
DR   SMART; SM00360; RRM; 4.
DR   PROSITE; PS50102; RRM; 4.
DR   PROSITE; PS00030; RRM_RNP_1; 3.
KW   Nuclear protein; Phosphorylation; Methylation; DNA-binding; Repeat;
KW   RNA-binding.
FT   INIT_MET      0      0
FT   DOMAIN      141    169       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      188    213       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      239    271       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      307    383       RNA-BINDING (RRM) 1.
FT   DOMAIN      393    466       RNA-BINDING (RRM) 2.
FT   DOMAIN      485    559       RNA-BINDING (RRM) 3.
FT   DOMAIN      571    646       RNA-BINDING (RRM) 4.
FT   DOMAIN      648    701       ARG/GLY/PHE-RICH.
FT   DOMAIN       56    133       8 X 8 AA TANDEM REPEATS OF X-T-P-X-K-K-
FT                                X-X.
FT   REPEAT       56     63       1.
FT   REPEAT       73     80       2.
FT   REPEAT       81     88       3.
FT   REPEAT       89     96       4.
FT   REPEAT       97    102       5 (INCOMPLETE).
FT   REPEAT      103    110       6.
FT   REPEAT      118    125       7.
FT   REPEAT      126    133       8.
FT   REPEAT      217    221       X-T-P-X-K-K-X-X MOTIF.
FT   MOD_RES     143    143       PHOSPHORYLATION.
FT   MOD_RES     156    156       PHOSPHORYLATION.
FT   MOD_RES     187    187       PHOSPHORYLATION.
FT   MOD_RES     655    655       METHYLATION (DI-).
FT   MOD_RES     659    659       METHYLATION (DI-).
FT   MOD_RES     665    665       METHYLATION (DI-).
FT   MOD_RES     669    669       METHYLATION (DI-).
FT   MOD_RES     673    673       METHYLATION (DI-).
FT   MOD_RES     679    679       METHYLATION (DI-).
FT   MOD_RES     681    681       METHYLATION (DI-).
FT   MOD_RES     687    687       METHYLATION (DI-).
FT   MOD_RES     691    691       METHYLATION (DI-).
FT   MOD_RES     694    694       METHYLATION (DI-).
SQ   SEQUENCE   713 AA;  76997 MW;  79DDCF724CED7DB4 CRC64;


  Query Match           13.0%;  Score 206;  DB 1;  Length 713;
  Best Local Similarity  23.2%;  Pred. No. 0.00017;
  Matches   66;  Conservative   31;  Mismatches   81;  Indels  106;  Gaps    8;

Qy      97 EQDDFTSSDDEDVPE----------------------AVPAPA---------------- 117
           |::| :| ::  :|:                        ||| ||
Db      34 EEEDDSSGEEVVIPQKKGKKATATPAKKVVVSQTKKVAVPTPAKKAAVTPGKKAAATPAK 93

Qy     118 -----PTAVTANGNAGAAVVKADTKPKAKPAEVKPAE-------EKPESDEEDESDDEDE 165
                 ||   |  ||   ||       |   | ||:         | |  :||| ||:||
Db      94 KAVTPAKAVATPGKKGATQAKALVATPGKKGAVTPAKGAKNGKNAKKEDSDEDEDDDDDE 153

Qy     166 SEEDDDSEKGMDVDEDD-------------------SDDDEEEDSEDEEEEETPKKPEP 205
           : |:| |   | :||:                    |:||||| |:|||:|     |
Db     154 DDSDEDEE---DEEEDEFEPPVVKGKQGKVAAAAPASEDEDEEEDEEEEEEDE-----EE 205

Qy     206 INKKRPNESVSKTPVSGKKAKPAAAPAST-------------------------PQKTK 239
           :    |::  ||  ||||     |                              :: |
Db     206 EDDSEEEEAMEITPAKGKKAPAKVVPVKAKNVAEEDDDDEEEDEDEEEDEEEEEDEEEEE 265

Qy     240 KKKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGGNNNKKPFN 283
           :::      | | |:  |      : |::  |   |: :  |||
Db     266 EEEEEEPVKPAPGKR-KKEMTKQKEVPEAKKQKVEGSESTTPFN 308


RESULT   8
NUCL_HUMAN
ID   NUCL_HUMAN     STANDARD;      PRT;   706 AA.
AC   P19338;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Nucleolin (Protein C23).
GN   NCL.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89290043; PubMed=2737305;
RA   Srivastava M., Fleming P.J., Pollard H.B., Burns A.L.;
RT   "Cloning and sequencing of the human nucleolin cDNA.";
RL   FEBS Lett. 250:99-105(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90368666; PubMed=2394707;
RA   Srivastava M., McBride O.W., Fleming P.J., Pollard H.B., Burns A.L.;
RT   "Genomic organization and chromosomal localization of the human
RT   nucleolin gene.";
RL   J. Biol. Chem. 265:14922-14931(1990).
CC   -!- FUNCTION: NUCLEOLIN IS THE MAJOR NUCLEOLAR PROTEIN OF GROWING
CC       EUKARYOTIC CELLS. IT IS FOUND ASSOCIATED WITH INTRANUCLEOLAR
CC       CHROMATIN AND PRERIBOSOMAL PARTICLES. IT INDUCES CHROMATIN
CC       DECONDENSATION BY BINDING TO HISTONE H1. IT IS THOUGHT TO PLAY A
CC       ROLE IN PRE-RRNA TRANSCRIPTION AND RIBOSOME ASSEMBLY.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
CC   -!- SIMILARITY: CONTAINS 4 RNA RECOGNITION MOTIFS (RRM).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M60858; AAA59954.1; -.
DR   PIR; A35804; A35804.
DR   PIR; S04631; S04631.
DR   HSSP; P09651; 1HA1.
DR   Aarhus/Ghent-2DPAGE; 1210; NEPHGE.
DR   MIM; 164035; -.
DR   InterPro; IPR000504; RRM.
DR   Pfam; PF00076; rrm; 4.
DR   SMART; SM00360; RRM; 4.
DR   PROSITE; PS50102; RRM; 4.
DR   PROSITE; PS00030; RRM_RNP_1; 3.
```

```
KW   Nuclear protein; Phosphorylation; Methylation; DNA-binding; Repeat;
KW   RNA-binding.
FT   INIT_MET      0      0
FT   DOMAIN      142    170       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      184    208       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      233    270       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      306    382       RNA-BINDING (RRM) 1.
FT   DOMAIN      392    465       RNA-BINDING (RRM) 2.
FT   DOMAIN      485    559       RNA-BINDING (RRM) 3.
FT   DOMAIN      571    643       RNA-BINDING (RRM) 4.
FT   DOMAIN      645    694       ARG/GLY/PHE-RICH.
FT   DOMAIN       57    134       8 X 8 AA TANDEM REPEATS OF X-T-P-X-K-K-
FT                               X-X.
FT   REPEAT       57     64       1.
FT   REPEAT       74     81       2.
FT   REPEAT       82     89       3.
FT   REPEAT       90     97       4.
FT   REPEAT       98    103       5 (INCOMPLETE).
FT   REPEAT      104    111       6.
FT   REPEAT      119    126       7.
FT   REPEAT      127    134       8.
FT   MOD_RES     144    144       PHOSPHORYLATION (BY SIMILARITY).
FT   MOD_RES     152    152       PHOSPHORYLATION (BY SIMILARITY).
FT   MOD_RES     183    183       PHOSPHORYLATION (BY SIMILARITY).
SQ   SEQUENCE   706 AA;  76213 MW;  85A2F2CA22EA03DB CRC64;
```

```
  Query Match             12.9%;  Score 205;  DB 1;  Length 706;
  Best Local Similarity   27.2%;  Pred. No. 0.00019;
  Matches   74;  Conservative   29;  Mismatches   95;  Indels   74;  Gaps    9;

Qy       6 VAVTP-KNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGG------AKLVIGTLS 58
           | |:| |     || : : |   : :     |   :   :|| |       | :: |
Db      63 VVVSPTKKVAVATPAKKAAVTPGKKAAATPAKKTVTPAKAVTTPGKKGATPGKALVATPG 122

Qy      59 QDKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDF--TSSDDEDVPEAVPAP 116
           :             |   :   | |   :     |   | ||      ||||  |
Db     123 K-------------KGAAIPAKGAKNGKNAKKEDSDEEEDDDSEEDEEDDEDEDEDEDEI 169

Qy     117 APTAVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEE-------- 168
             | |: |   | ||          ||     :|  | ||:|: |:||:|||
Db     170 EPAAMKA---AAAA-----------PASEDEDDEDDEDDEDDDDDEEDDSEEEAMETTPA 215

Qy     169 -------------------DDDSEKGMDVDEDDSDDDEEEDSEDEEEEETPKKPEPINKK 209
                             :|: |   | |:|| :|||::| ||::|||  :: ||
Db     216 KGKKAAKVVPVKAKNVAEDEDEEEDDEDEDDDDDEDEDDDDEDDEEEEEEEEEEP---- 271

Qy     210 RPNESVSKTPVSGKK--AKPAAAPASTPQKTE 239
               | : |   ||  ||  ||| :  || |
Db     272 -----VKEAPGKRKKEMAKQKAAPEAKKQKVE 298
```

```
RESULT   9
HS32_DICDI
ID   HS32_DICDI      STANDARD;      PRT;   281 AA.
AC   P54658;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   32 kDa heat shock protein (4-1 protein).
GN   HSPC OR HSP32.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyostelilda; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96001265; PubMed=7557471;
RA   de Maria A.C., Gomes S.L., Juliania M.H., Mazzarella R., Klein C.;
RT   "Cloning of a cDNA encoding a novel heat-shock protein from
RT   Dictyostelium discoideum.";
RL   Gene 163:163-164(1995).
CC   -!- DEVELOPMENTAL STAGE: PRESENT AT HIGH LEVELS IN GROWING CELLS BUT
CC       DECREASES DRAMATICALLY DURING THE EARLY HOURS OF DEVELOPMENT.
CC   -!- INDUCTION: BY HEAT SHOCK.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U40211; AAC47710.1; -.
DR   EMBL; L39778; AAA99510.1; -.
DR   DictyDb; DD01076; hspC.
KW   Heat shock.
FT   DOMAIN      251    257       POLY-GLN.
FT   DOMAIN      262    265       POLY-ASN.
SQ   SEQUENCE   281 AA;  31481 MW;  2C64F57C847D7CA4 CRC64;
```

```
  Query Match             12.6%;  Score 200.5;  DB 1;  Length 281;
  Best Local Similarity   21.6%;  Pred. No. 0.00014;
  Matches   72;  Conservative   49;  Mismatches  123;  Indels   89;  Gaps   10;

Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQD 60
           |:|:|  || :    :  :|  : |:::| :    :    | |:        ||  : :|
Db       2 MQFFGTIVTKEEPVNLELDEGDIFHLTKAIIHPKSQGKGKVYLTA-------VISLMEED 54

Qy      61 KFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTA 120
           :                                :|:||    |:|: |       |
Db      55 E---------------------------------MEEDDV--DDEEESPREDIVEIPIG 78

Qy     121 VTANG-----------NAGAAVVKADTKPKAKPAEVKPAE------EKPESDEEDESDDE 163
           :   |           | |  :|: : : :     |          |:   |:||  |:
Db      79 ILEAGKIDQIDLNLHYNFG-QIVRFELQAENGAGYVVALSGSVITMEQGGCDDEDCDDEH 137

Qy     164 DESEEDDDSEKGMDVDEDDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGK 223
             : |||:     ::| |:   | ::| || ::|| |:   |  ||     :|: | | |
Db     138 CINHEDDE-----EIDSDEEFGDSDQDEEDSDDEEIPQLIAPATKKGKITEISEVPESKK 192

Qy     224 KAKPAAAPASTPQKTE----------------KKKGGHTATPHPAKKGGKSPVNANQSP 266
           :  |       |:| :                 :|     |   |||       || : |
Db     193 EKTPEPKKVPEPKKEQVKQPTQPQQKKAAAQQPEKANNKPAAASPAKPQNNQSKNAPKQP 252

Qy     267 KSGGQS-SGGNNNKKPFNSGKQFGGSNNKGSNK 298
           :   || :  ||||:| |  :     |||   |
Db     253 QQQQQSPAKNNNNKRPQNQNE-----NNKKKQK 280
```

```
RESULT  10
NP14_RAT
ID   NP14_RAT       STANDARD;      PRT;   704 AA.
AC   P41777;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Nucleolar phosphoprotein p130 (Nucleolar 130 kDa protein) (140 kDa
DE   nucleolar phosphoprotein) (Nopp140) (Nucleolar and coiled-body
DE   phosphoprotein 1).
GN   NOLC1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 292-309 AND 563-601.
RC   TISSUE=Liver;
RX   MEDLINE=92323542; PubMed=1623516;
RA   Meier U.T., Blobel G.;
RT   "Nopp140 shuttles on tracks between nucleolus and cytoplasm.";
RL   Cell 70:127-138(1992).
RN   [2]
RP   INTERACTION WITH NOP5 AND FIBRILLARIN.
RX   MEDLINE=20143579; PubMed=10679015;
```

```
RA   Yang Y., Isaac C., Wang C., Dragon F., Pogacic V., Meier U.T.;
RT   "Conserved composition of mammalian box H/ACA and box C/D small
RT   nucleolar ribonucleoprotein particles and their interaction with the
RT   common factor Nopp140.";
RL   Mol. Biol. Cell 11:567-577(2000).
CC   -!- FUNCTION: RELATED TO NUCLEOLOGENESIS, MAY PLAY A ROLE IN THE
CC       MAINTENANCE OF THE FUNDAMENTAL STRUCTURE OF THE FIBRILLAR CENTER
CC       AND DENSE FIBRILLAR COMPONENT IN THE NUCLEOLUS. IT HAS INTRINSIC
CC       GTPASE AND ATPASE ACTIVITIES. MAY PLAY AN IMPORTANT ROLE IN
CC       TRANSCRIPTION CATALYZED BY RNA POLYMERASE I (BY SIMILARITY).
CC   -!- SUBUNIT: Interacts with Dkc1/Nap57, Nop5/Nap65 and fibrillarin.
CC   -!- SUBCELLULAR LOCATION: SHUTTLES ON CURVILINEAR TRACKS BETWEEN
CC       NUCLEOLUS AND CYTOPLASM. THESE TRACKS EXTEND FROM THE DENSE
CC       FIBRILLAR COMPONENT OF THE NUCLEOLUS ACROSS THE NUCLEOPLASM TO
CC       A LIMITED NUMBER OF NUCLEAR PORE COMPLEXES.
CC   -!- PTM: THIS PROTEIN UNDERGOES RAPID AND MASSIVE PHOSPHORYLATION AND
CC       DEPHOSPHORYLATION ON CK-II AND PKC SITES. NOPP140 IS ONE OF THE
CC       MOST PHOSPHORYLATED PROTEINS IN THE CELL.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M94287; AAA41718.1; -.
DR   EMBL; M94288; AAA41719.1; -.
KW   Nuclear protein; Phosphorylation; Repeat; GTP-binding; ATP-binding.
FT   DOMAIN       84    570       11 X 12 AA APPROXIMATE REPEATS OF AN
FT                                ACIDIC SERINE CLUSTER.
FT   REPEAT        84     95       ACIDIC SERINE CLUSTER 1.
FT   REPEAT       127    138       ACIDIC SERINE CLUSTER 2.
FT   REPEAT       170    181       ACIDIC SERINE CLUSTER 3.
FT   REPEAT       231    242       ACIDIC SERINE CLUSTER 4.
FT   REPEAT       274    285       ACIDIC SERINE CLUSTER 5.
FT   REPEAT       335    346       ACIDIC SERINE CLUSTER 6.
FT   REPEAT       373    384       ACIDIC SERINE CLUSTER 7.
FT   REPEAT       434    445       ACIDIC SERINE CLUSTER 8.
FT   REPEAT       479    490       ACIDIC SERINE CLUSTER 9.
FT   REPEAT       524    535       ACIDIC SERINE CLUSTER 10.
FT   REPEAT       559    570       ACIDIC SERINE CLUSTER 11.
FT   MOD_RES      567    567       PHOSPHORYLATION (BY CK2).
FT   VARIANT      150    150       MISSING (IN NOPP140B).
SQ   SEQUENCE   704 AA;  73562 MW;  14DF1BF2DE483EA3 CRC64;


  Query Match           11.8%;  Score 188;  DB 1;  Length 704;
  Best Local Similarity  26.6%;  Pred. No. 0.0015;
  Matches   78;  Conservative   38;  Mismatches   85;  Indels   92;  Gaps   15;

Qy      89 IGYKSP--------------NIEQDDFTSSDDEDVP-----EAVPA-PAPTAVTANGN- 126
           :| :||                : |: | :| :::  |        || |||    |  :
Db     314 VGAQSPKKAAAQTQPADSSADSSEESDSSSEEEKKTPAKTVVSKTPAKPAPVKKKAESSS 373

Qy     127 -----------AGAAVVKADTKPKAKPAEV-KP-------------AEEKPESDEEDES 160
                     | |  | |   | :|||   ||              : :||:| : | |
Db     374 DSSDSDSSEDEAPAKPVSATKSPLSKPAVTPKPPAAKAVATPKQPAGSGQKPQSRKADSS 433

Qy     161 DDEDE--SEEDDDSEKGMDVDE-------------------DDSDDDEEEDSEDEEEEE 198
             |:|  | |:: ::| :  :                   ||  |  | |  |||::
Db     434 SSEEESSSSEEEEATKKSVTTPKARVTAKAAPSLPAKQAPRAGGDSSSD-SESSSSEEEKK 492

Qy     199 TPKKPEPINKKRPNESVSK-TPVSGKKAKPAAAPASTPQKT--EKKKGGHTATPHPAKKG 255
           || || |  ||    :| | |||    |: ::: :|:   :  ||||    ||| | : |
Db     493 TPPKP-PAKKKAAGAAVPKPTPVKKAAAESSSSSSSSEDSSEEEKKKPKSKATPKP-QAG 550

Qy     256 GKSPVNANQSPKSGGQS----SGGNNNKKPFNSGKQFGGSNNKGSNKGKGKGR 304
             : | |:|: |:| :|         |||              |:  | || |
Db     551 KANGVPASQNGKAGKESEEEEEDTEQNKKA-----------AGTKPGSGKKR 591
```

```
RESULT  11
SSRP_DROME
ID   SSRP_DROME     STANDARD;      PRT;   723 AA.
AC   Q05344; Q9W1J4;
DT   01-OCT-1994 (Rel. 30, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Single-strand recognition protein (SSRP) (Chorion-factor 5).
GN   SSRP OR SSRP1 OR CF5 OR CG4817.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=93342017; PubMed=7688122;
RA   Hsu T., King D.L., Labonne C., Kafatos F.C.;
RT   "A Drosophila single-strand DNA/RNA-binding factor contains a high-
RT   mobility-group box and is enriched in the nucleolus.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:6488-6492(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93241947; PubMed=8479916;
RA   Bruhn S.L., Housman D.E., Lippard S.J.;
RT   "Isolation and characterization of cDNA clones encoding the
RT   Drosophila homolog of the HMG-box SSRP family that recognizes
RT   specific DNA structures.";
RL   Nucleic Acids Res. 21:1643-1646(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- FUNCTION: BINDS TO SINGLE-STRANDED DNA AND RNA SEQUENCES, WITH
```

```
CC       HIGHEST AFFINITY FOR NUCLEOTIDES G AND U.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR. PARTICULARLY ABUNDANT IN THE
CC       NUCLEOLUS.
CC   -!- TISSUE SPECIFICITY: EXPRESSED MOST ABUNDANTLY IN NURSE CELLS IN
CC       THE OVARY.
CC   -!- DEVELOPMENTAL STAGE: ABUNDANT THROUGHOUT OOGENESIS AND
CC       EMBRYOGENESIS, DECREASES DURING LARVAL STAGES AND INCREASES
CC       AGAIN IN PUPAE.
CC   -!- SIMILARITY: CONTAINS 1 HMG BOX.
CC   -!- SIMILARITY: BELONGS TO THE SSRP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L08825; AAA28914.1; -.
DR   EMBL; X68408; CAA48471.1; -.
DR   EMBL; AE003462; AAF47064.1; -.
DR   PIR; A48217; A48217.
DR   HSSP; Q05783; 1HMA.
DR   TRANSFAC; T01848; -.
DR   FlyBase; FBgn0010278; Ssrp.
DR   InterPro; IPR000910; HMG_12_box.
DR   InterPro; IPR000969; SSrcognition.
DR   Pfam; PF00505; HMG_box; 1.
DR   PRINTS; PR00887; SSRCOGNITION.
DR   SMART; SM00398; HMG; 1.
KW   DNA-binding; RNA-binding; Nuclear protein.
FT   DNA_BIND     555     621       HMG BOX.
FT   CONFLICT      13      13       E -> Q (IN REF. 1).
FT   CONFLICT      33      33       K -> E (IN REF. 1).
FT   CONFLICT     212     212       G -> R (IN REF. 1).
FT   CONFLICT     428     428       M -> T (IN REF. 1).
FT   CONFLICT     498     498       V -> E (IN REF. 1).
FT   CONFLICT     504     504       D -> E (IN REF. 1).
FT   CONFLICT     573     573       I -> Y (IN REF. 1).
SQ   SEQUENCE   723 AA;  81532 MW;  DE8017F75C6CA207 CRC64;
```

```
  Query Match             11.8%;  Score 188;  DB 1;  Length 723;
  Best Local Similarity   23.5%;  Pred. No. 0.0015;
  Matches   76;  Conservative   53;  Mismatches  110;  Indels   84;  Gaps   14;

Qy      19 EEDSLVHISQASLDCTVKSGESVVLSVTV-GGAKLVIGTLSQDKFPQISFDLVFDKEFEL 77
           || | |: :::       | |    ||:  |   :  :: :::: :: || :  |:  :
Db     372 EEISSVNFARSG-----GSTRSFDFEVTLKNGTVHIFSSIEKEEYAKL-FDYITQKKLHV 425

Qy      78 SHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTAVTANGNAGAAVVKADTK 137
           |: |   :    |||       ||  ||:|: |:|                | :||: :
Db     426 SNMGKDKS----GYKDV-----DFGDSDNENEPDAY--------------LARLKAEAR 461

Qy     138 PKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMDVD---EDDSDDDE------- 187
           |         |:  : | ::|| |||    :::|:   : |   | |||||
Db     462 EK--------EEDDDDGDSDEESTDEDFKPNENESDVAEEYDSNVESDSDDDSDASGGGG 513

Qy     188 EEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKK--AKPAAAPAS-------TPQKT 238
           : |   :::|:  :| |   ||   :  :| |   ||   ||  | :       | :
Db     514 DSDGAKKKKEKKSEKKEKKEKKHKEKERTKKPSKKKKDSGKPKRATTAFMLWLNDTRESI 573

Qy     239 EKKKGGHTATPHPAKKGGK----------------------SPVNANQSPKSGGQS---S 273
           ::: |   |   |||||:                      |  |::|| |
Db     574 KRENPGIKVT-EIAKKGGEMWKELKDKSKWEDAAAKDKQRYHDEMRNYKPEAGGDSDNEK 632

Qy     274 GGNNNKKPFNSGKQFGGSNNKGS 296
           || ::||          :|  ||
Db     633 GGKSSKKRKTEPSPSKKANTSGS 655


RESULT  12
```

```
NPM_CHICK
ID   NPM_CHICK      STANDARD;      PRT;   294 AA.
AC   P16039;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Nucleophosmin (NPM) (Nucleolar phosphoprotein B23) (Numatrin)
DE   (Nucleolar protein NO38).
GN   NPM1.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90206792; PubMed=2320420;
RA   Maridor G., Nigg E.A.;
RT   "cDNA sequences of chicken nucleolin/C23 and NO38/B23, two major
RT   nucleolar proteins.";
RL   Nucleic Acids Res. 18:1286-1286(1990).
RN   [2]
RP   DISCUSSION OF SEQUENCE.
RX   MEDLINE=90304215; PubMed=2114180;
RA   Maridor G., Krek W., Nigg E.A.;
RT   "Structure and developmental expression of chicken nucleolin and
RT   NO38: coordinate expression of two abundant non-ribosomal nucleolar
RT   proteins.";
RL   Biochim. Biophys. Acta 1049:126-133(1990).
CC   -!- FUNCTION: ASSOCIATED WITH NUCLEOLAR RIBONUCLEOPROTEIN STRUCTURES
CC       AND BIND SINGLE-STRANDED NUCLEIC ACIDS. IT MAY FUNCTION IN THE
CC       ASSEMBLY AND/OR TRANSPORT OF RIBOSOME.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR. GENERALLY NUCLEOLAR, BUT IS
CC       TRANSLOCATED TO THE NUCLEOPLASM IN CASE OF SERUM STARVATION OR
CC       TREATMENT WITH ANTICANCER DRUGS (BY SIMILARITY).
CC   -!- PTM: PHOSPHORYLATED.
CC   -!- SIMILARITY: BELONGS TO THE NUCLEOPLASMIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X17200; CAA35061.1; -.
DR   PIR; S08415; DNCHFM.
DR   InterPro; IPR004301; Nucleoplasmin.
DR   Pfam; PF03066; Nucleoplasmin; 1.
KW   Nuclear protein; Phosphorylation; RNA-binding.
FT   DOMAIN        1      11       MET-RICH.
FT   DOMAIN      122     133       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      153     158       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   DOMAIN      161     186       ASP/GLU-RICH (HIGHLY ACIDIC).
FT   DOMAIN      190     196       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ   SEQUENCE   294 AA;  32632 MW;  B1FF89B5F2322DED CRC64;
```

```
  Query Match             11.5%;  Score 183.5;  DB 1;  Length 294;
  Best Local Similarity   28.7%;  Pred. No. 0.0012;
  Matches   73;  Conservative   32;  Mismatches   86;  Indels   63;  Gaps   13;

Qy      15 KVTPEEDSLVHISQASL-DCTVKSG--------ESVVLSVTVGGAKLVIGTLSQDKFPQI 65
           || |||:       | ||   |: :|        |:  |        |:|: :|     | :
Db      34 KVDDEENE----HQLSLRTVTLGAGAKDELHVVEAEALDYEGNPTKVVLASLKMSVQPTV 89

Qy      66 S---FDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTAVT 122
           |   |::     | |        |:  |     :|::  :  ::||             |
Db      90 SLGGFEIT--PPFVLRLKCGSGPVYVSGQHLVALEEEPESEDEEED------------T 134

Qy     123 ANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMDVDEDD 182
             ||       | ||  |     |  ::||:  |:|| ||||| :::|        ||||
```

```
Db    135 KIGN-------ASTKRPASGGGAKTPQKKPKLSEDDEDDDEDEDDDED--------DEDD 179

Qy    183 SDDDEEEDSEDEEEEETP-KKP--EPINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTE 239
          ||       |||| :|| ||| || |     :: |  :|| :||:   : |  ||
Db    180 LDD-------DEEEIKTPMKKPAREPAGK-----NMQKAKQNGKDSKPSTPASKT--KTP 225

Qy    240 KKKGGHTATPHPAK 253
          |   : ||   |
Db    226 DSKKDKSLTPKTPK 239


RESULT 13
T2FA_HUMAN
ID   T2FA_HUMAN      STANDARD;      PRT;   517 AA.
AC   P35269;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Transcription initiation factor IIF, alpha subunit (TFIIF-alpha)
DE   (Transcription initiation factor RAP74).
GN   GTF2F1 OR RAP74.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=92131135; PubMed=1734283;
RA   Aso T., Vasavada H.A., Kawaguchi T., Germino F.J., Ganguly S.,
RA   Kitajima S., Weissman S.M., Yasukochi Y.;
RT   "Characterization of cDNA for the large subunit of the transcription
RT   initiation factor TFIIF.";
RL   Nature 355:461-464(1992).
RN   [2]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=92131136; PubMed=1734284;
RA   Finkelstein A., Kostrub C.F., Li J., Chavez D.P., Wang B.Q.,
RA   Fang S.M., Greenblatt J., Burton Z.F.;
RT   "A cDNA encoding RAP74, a general initiation factor for transcription
RT   by RNA polymerase II.";
RL   Nature 355:464-467(1992).
CC   -!- FUNCTION: TFIIF IS A GENERAL TRANSCRIPTION INITIATION FACTOR THAT
CC       BINDS TO RNA POLYMERASE II AND HELPS TO RECRUIT IT TO THE
CC       INITIATION COMPLEX IN COLLABORATION WITH TFIIB. IT PROMOTES
CC       TRANSCRIPTION ELONGATION.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- PTM: PHOSPHORYLATED.
CC   -!- SIMILARITY: TO OTHER TRANSCRIPTION FACTOR IIF, ALPHA SUBUNIT.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X64037; CAA45408.1; -.
DR   EMBL; X64002; CAA45404.1; -.
DR   PIR; S20248; S20248.
DR   TRANSFAC; T02168; -.
DR   MIM; 189968; -.
KW   Transcription regulation; DNA-binding; Nuclear protein;
KW   Phosphorylation.
FT   CONFLICT    231    231       V -> I (IN REF. 2).
FT   CONFLICT    361    361       F -> L (IN REF. 2).
SQ   SEQUENCE   517 AA;  58274 MW;  F0D2BE44D2F3820F CRC64;


  Query Match             11.2%;  Score 178.5;  DB 1;  Length 517;
  Best Local Similarity   26.8%;  Pred. No. 0.0036;
  Matches   74;  Conservative   34;  Mismatches   83;  Indels   85;  Gaps   13;

Qy     72 DKEPELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTAVTANGNAGAAV 131
          :||        ::  :|       ::| |   |||  |              |:|  |  |
Db    193 EKEKRGRRKASELRIH-------DLEDDLEMSSDASD-------------ASGEEGGRV 231

Qy    132 VKADTK-PKAKPAEVKPAEEKPESDEE--DESDDED-------------ESEEDDDSE- 173
          ||  | | ||    |  ::|  ||:|  ::||| |               |:|: :|:
Db    232 PKAKKKAPLAKGGRKK--KKKKGSDDEAPEDSDDGDFEGQEVDYMSDGSSSSQEEPESKA 289

Qy    174 ---------KGMDVDEDDSDDDEEE--DSEDEEEEETPKKPEPINKKRPNESVSKTPVS- 221
                  ||:|   | |:: |||    ||:|||| | | |  |||  :|  ::  |
Db    290 KAPQQEEGPKGVDEQSDSSEESEEEKPPEEDKEEEEEKKAPTPQEKKRRKDSSEESDSSE 349

Qy    222 ----GKKAKPA--AAPASTPQKTEKKKGGHTATPH-----PAKKGGKS------------ 258
              :|  |   |   || | |:|  | ::  :     |: :|| :
Db    350 ESDIDSEASSAFFMAKKKTPPKRERKPSGGSSRGNSRPGTPSAEGGSTSSTLRAAASKLE 409

Qy    259 --------PVNANQSPKSGGQSSGGNNNKKPFNSGK 286
                  |       :| ||  | :  :|  |||
Db    410 QGKRVSEMPAAKRLRLDTGPQSLSGKSTPQP-PSGK 444


RESULT 14
NFM_RABIT
ID   NFM_RABIT      STANDARD;      PRT;   644 AA.
AC   P54938;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Neurofilament triplet M protein (160 kDa neurofilament protein)
DE   (Neurofilament medium polypeptide) (NF-M) (Fragment).
GN   NEF3 OR NEFM OR NFM.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Heart;
RX   MEDLINE=97055255; PubMed=8899542;
RA   Vitadello M., Vettore S., Lamar E., Chien K.R., Gorza L.;
RT   "Neurofilament M mRNA is expressed in conduction system myocytes of
RT   the developing and adult rabbit heart.";
RL   J. Mol. Cell. Cardiol. 28:1833-1844(1996).
CC   -!- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
CC       AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
CC   -!- PTM: THERE ARE A NUMBER OF REPEATS OF THE TRIPEPTIDE K-S-P, NFM IS
CC       PHOSPHORYLATED ON A NUMBER OF THE SERINES IN THIS MOTIF. IT IS
CC       THOUGHT THAT PHOSPHORYLATION OF NFH RESULTS IN THE FORMATION OF
CC       INTERFILAMENT CROSS BRIDGES THAT ARE IMPORTANT IN THE MAINTENANCE
CC       OF AXONAL CALIBER.
CC   -!- PTM: PHOSPHORYLATION SEEMS TO PLAY A MAJOR ROLE IN THE FUNCTIONING
CC       OF THE LARGER NEUROFILAMENT POLYPEPTIDES (NF-M AND NF-H), THE
CC       LEVELS OF PHOSPHORYLATION BEING ALTERED DEVELOPMENTALLY AND
CC       COINCIDENT WITH A CHANGE IN THE NEUROFILAMENT FUNCTION.
CC   -!- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z47378; CAA87454.1; -.
DR   InterPro; IPR001664; IF.
DR   Pfam; PF00038; filament; 1.
DR   PROSITE; PS00226; IF; 1.
KW   Intermediate filament; Coiled coil; Neurone; Phosphorylation;
KW   Glycoprotein.
FT   NON_TER       1      1
FT   DOMAIN       <1    197       ROD.
```

```
FT   DOMAIN      198    643       TAIL.
FT   DOMAIN       <1     33       COIL 1B.
FT   DOMAIN       34     50       LINKER 12.
FT   DOMAIN       51     72       COIL 2A.
FT   DOMAIN       73     76       LINKER 2.
FT   DOMAIN       77    197       COIL 2B.
FT   CARBOHYD    217    217       O-LINKED (GLCNAC) (BY SIMILARITY).
SQ   SEQUENCE   644 AA;  72450 MW;  030FDAA622889678 CRC64;


  Query Match             11.1%;  Score 177;  DB 1;  Length 644;
  Best Local Similarity   23.3%;  Pred. No. 0.0053;
  Matches   76;  Conservative   49;  Mismatches  125;  Indels   76;  Gaps   12;

Qy      40 SVVLSVTVGGAKLVIGTLS-QDKFPQ--ISFDLVFDKEFELSHSGTK-ANVHFIGYKSPN 95
           || :|  :   |:    |  | || :  |    | |:: |:  : |  |    :  |
Db     215 SVTISSKIQKTKVEAPKLKVQHKFVEEIIEETKVEDEKSEMEDALTAIAEELAVSVKEEE 274

Qy      96 IEQDDFTSSDDEDVPEAVPAPAPTAVTANGNAGAAVVKADTKPKAKPAEVKPAEE--KPE 153
           |::     ::::  | | |   : | |           | |: |  | :  ||  : |
Db     275 KEEEAEGKEEEQEAEEEVAAAKKSPVKA-----------TTPEIKEEEGEKEEEGQEEE 322

Qy     154 SDEEDESDDEDESEEDDDSEKGMDVDEDDSDDDE-------------EEDSEDEEEEETP 200
           :||||    |::||    ::|   :| : :: |                ||   :|: ||
Db     323 EEEEDEGVKSDQAEEGGSEKEGSSKNEGEQEEGETEAEGEVEEAEAKEEKKTEEKSEEVA 382

Qy     201 KKPEPINK------KRPNESVSKTPV----------SGK-----------KAKPAAAPA 232
           | ||: :      ::    | |:||          :||           : | ||  :
Db     383 AKEEPVTEAKVGKPEKAKSPVPKSPVEEVKPKAEATAGKGEQKEEEEKVEEEKKKAAKES 442

Qy     233 STPQKTEKKKGGHTATPHPAKKGGKSPVNANQS--------PKSGGQSSGGNNNKKPFNS 284
             :| |||:      |   ||  :|||    :        |   |         :||
Db     443 PKEEKVEKKEEKPKDVP---KKKAESPVKEEAAEEAATITKPTKVGLEKETKEGEKPLQQ 499

Qy     285 GK------QFGGSNNKGSNKGKGKGR 304
           |       : |||  :||::|  : :
Db     500 EKEKEKAGEEGGSEEEGSDQGSKRAK 525


RESULT  15
NUCL_XENLA
ID   NUCL_XENLA     STANDARD;      PRT;   650 AA.
AC   P20397;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Nucleolin (Protein C23).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=93181171; PubMed=8441611;
RA   Rankin M.L., Heine M.A., Xiao S., Leblanc M.D., Nelson J.W.,
RA   Dimario P.J.;
RT   "A complete nucleolin cDNA sequence from Xenopus laevis.";
RL   Nucleic Acids Res. 21:169-169(1993).
RN   [2]
RP   SEQUENCE OF 125-650 FROM N.A.
RX   MEDLINE=89252811; PubMed=2656405;
RA   Caizergues-Ferrer M., Mariottini P., Curie C., Lapeyre B., Gas N.,
RA   Amalric F., Amaldi F.;
RT   "Nucleolin from Xenopus laevis: cDNA cloning and expression during
RT   development.";
RL   Genes Dev. 3:324-333(1989).
CC   -!- FUNCTION: NUCLEOLIN IS THE MAJOR NUCLEOLAR PROTEIN OF GROWING
CC       EUKARYOTIC CELLS. IT IS FOUND ASSOCIATED WITH INTRANUCLEOLAR
CC       CHROMATIN AND PRERIBOSOMAL PARTICLES. IT INDUCES CHROMATIN
CC       DECONDENSATION BY BINDING TO HISTONE H1. IT IS THOUGHT TO PLAY A
CC       ROLE IN PRE-RRNA TRANSCRIPTION AND RIBOSOME ASSEMBLY.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
CC   -!- SIMILARITY: CONTAINS 4 RNA RECOGNITION MOTIFS (RRM).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X63091; CAA44805.1; -.
DR   PIR; A30166; A30166.
DR   PIR; S18874; S18874.
DR   PIR; S30250; S30250.
DR   HSSP; P09651; 1HA1.
DR   InterPro; IPR000504; RRM.
DR   Pfam; PF00076; rrm; 4.
DR   SMART; SM00360; RRM; 4.
DR   PROSITE; PS50102; RRM; 4.
DR   PROSITE; PS00030; RRM_RNP_1; 3.
KW   Nuclear protein; Phosphorylation; Methylation; DNA-binding; Repeat;
KW   RNA-binding.
FT   INIT_MET      0      0       BY SIMILARITY.
FT   DOMAIN      119    134       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      155    165       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      183    202       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      232    308       RNA-BINDING (RRM) 1.
FT   DOMAIN      324    398       RNA-BINDING (RRM) 2.
FT   DOMAIN      414    487       RNA-BINDING (RRM) 3.
FT   DOMAIN      502    577       RNA-BINDING (RRM) 4.
FT   DOMAIN      583    643       ARG/GLY/PHE-RICH.
FT   MOD_RES     154    154       PHOSPHORYLATION (BY SIMILARITY).
FT   CONFLICT    214    214       P -> Q (IN REF. 2).
FT   CONFLICT    218    219       PE -> LR (IN REF. 2).
FT   CONFLICT    410    410       E -> Q (IN REF. 2).
FT   CONFLICT    580    580       D -> E (IN REF. 2).
SQ   SEQUENCE   650 AA;  70064 MW;  5BF9BF09768E71B4 CRC64;


  Query Match             11.1%;  Score 177;  DB 1;  Length 650;
  Best Local Similarity   27.4%;  Pred. No. 0.0053;
  Matches   55;  Conservative   27;  Mismatches   45;  Indels   74;  Gaps    8;

Qy      95 NIEQDDFTSSDDE-DVP--------EAVPAPA-PTAVTANGNAGAAVVK---------- 133
           ::|:||  |||:| :||          | || | |       |  ||   |
Db      32 DMEEDD--SSDEEVEVPVKKTPAKKTATPAKATPGKAATPGKKGATPAKNGKQAKKQESE 89

Qy     134 --------------ADTKPKAKPAEVKPAEEKPESDEEDESDDE--------------- 163
                         || || |  |  | : : |:|||||::|
Db      90 EEEDDSDEEAEDQKPIKNKPVAKKAVAKKEESEEDDDDEDESEEEKAVAKKPTPAKKPAG 149

Qy     164 --DESEEDDDSE---KGMDV---------------DEDDSDDDEEEDSEDEEEEETPKK 202
             ||||:|| |   : |:|                ||:: |||||:| ::||:: : |:
Db     150 KKQESEEEDDEESEDEPMEVAPALKGKKTAQAAEEDDEEEDDDDEEDDDDEEEQQGSAKR 209

Qy     203 PEPINKKRPNESVSKTPVSGK 223
           : : |  |    :||  : :
Db     210 KKEMPKTIPEAKKTKTDTASE 230


Search completed: May  2, 2002, 22:31:10
Job time: 1416 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: May 2, 2002, 22:30:22 ; Search time 113.15 Seconds
(without alignments)
466.314 Million cell updates/sec

Title: US-09-645-337-8
Perfect score: 1589
Sequence: 1 MEFWGVAVTPKNATKVTPEE.....KQFGGSNNKGSNKGKGKGRA 305

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1589	100.0	305	10	O49209	O49209 arabidopsis
2	1578.5	99.3	306	10	Q9FNJ6	Q9fnj6 arabidopsis
3	594	37.4	245	10	Q9FVE6	Q9fve6 arabidopsis
4	594	37.4	257	10	O22238	O22238 arabidopsis
5	574	36.1	297	10	Q9M4T5	Q9m4t5 oryza sativ
6	543.5	34.2	307	10	O24591	O24591 zea mays (m
7	541	34.0	303	10	Q9M4U5	Q9m4u5 zea mays (m
8	532	33.5	300	10	Q9M4U4	Q9m4u4 zea mays (m
9	453.5	28.5	294	10	Q9M4T4	Q9m4t4 arabidopsis
10	450.5	28.4	296	10	Q9LZR5	Q9lzr5 arabidopsis
11	364	22.9	286	10	Q94F81	Q94f81 zea mays (m
12	262.5	16.5	203	10	Q9M4T3	Q9m4t3 arabidopsis
13	258	16.2	312	5	Q26711	Q26711 trypanosoma
14	230	14.5	798	5	O45181	O45181 caenorhabdi
15	210.5	13.2	346	5	Q9NLA3	Q9nla3 asterina pe
16	207.5	13.1	611	10	Q41042	Q41042 pisum sativ
17	205.5	12.9	707	11	Q99K50	Q99k50 mus musculu
18	204.5	12.9	715	11	Q9QZX1	Q9qzx1 rattus norv
19	200.5	12.6	570	5	Q9VZR7	Q9vzr7 drosophila
20	198.5	12.5	444	11	Q9CT46	Q9ct46 mus musculu
21	197	12.4	2163	5	Q9NFB6	Q9nfb6 plasmodium
22	196.5	12.4	635	10	Q40363	Q40363 medicago sa
23	194.5	12.2	688	5	Q9VNX6	Q9vnx6 drosophila
24	192.5	12.1	971	5	Q9XVS4	Q9xvs4 caenorhabdi
25	192	12.1	513	5	Q95PU8	Q95pu8 chironomus
26	189	11.9	857	5	Q9V6C4	Q9v6c4 drosophila
27	185	11.6	1089	12	O40947	O40947 kaposi's sa
28	185	11.6	1151	5	Q9XUR2	Q9xur2 caenorhabdi
29	184.5	11.6	759	10	Q9FM72	Q9fm72 arabidopsis
30	184.5	11.6	1129	12	Q9QR71	Q9qr71 kaposi's sa
31	184	11.6	416	12	Q9YPA9	Q9ypa9 kaposi's sa
32	183	11.5	976	12	Q9DUN0	Q9dun0 kaposi's sa
33	182.5	11.5	872	5	Q93178	Q93178 caenorhabdi
34	182.5	11.5	1187	11	P70475	P70475 rattus norv
35	182	11.5	337	13	Q91708	Q91708 xenopus lae
36	181	11.4	1036	12	Q9DUM3	Q9dum3 kaposi's sa
37	180.5	11.4	374	5	Q19973	Q19973 caenorhabdi
38	180	11.3	1003	12	Q91LX9	Q91lx9 kaposi's sa
39	180	11.3	1162	12	Q98148	Q98148 kaposi's sa
40	180	11.3	1300	12	O36421	O36421 alcelaphine
41	179	11.3	214	5	Q9N328	Q9n328 caenorhabdi
42	179	11.3	1331	11	Q9EPQ2	Q9epq2 mus musculu
43	178.5	11.2	477	10	Q93ZG9	Q93zg9 arabidopsis
44	178.5	11.2	517	4	Q9BWN0	Q9bwn0 homo sapien
45	178	11.2	1369	13	Q90XA6	Q90xa6 brachydanio

ALIGNMENTS

RESULT 1
O49209
ID O49209 PRELIMINARY; PRT; 305 AA.
AC O49209;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PUTATIVE HISTONE DEACETYLASE.
GN HD2 OR HD2B.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Dangl M., Haas H., Loidl P.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Wu K., Tian L., Malik K., Brown D., Miki B.;
RT "Functional analysis of HD2 histone deacetylase homologs in
RT Arabidopsis thaliana.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF044914; AAC02539.1; -.
DR EMBL; AF195546; AAG28473.1; -.
SQ SEQUENCE 305 AA; 32219 MW; 58AA6D5764930944 CRC64;

Query Match 100.0%; Score 1589; DB 10; Length 305;
Best Local Similarity 100.0%; Pred. No. 1.8e-98;
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQD 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQD 60

```
Qy      61 KFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 KFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTA 120

Qy     121 VTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMDVDE 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 VTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMDVDE 180

Qy     181 DDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTEK 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 DDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTEK 240

Qy     241 KKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGGNNNKKPFNSGKQFGGSNNKGSNKGK 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 KKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGGNNNKKPFNSGKQFGGSNNKGSNKGK 300

Qy     301 GKGRA 305
           |||||
Db     301 GKGRA 305
```

RESULT 2
Q9FNJ6
ID Q9FNJ6 PRELIMINARY; PRT; 306 AA.
AC Q9FNJ6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HISTONE DEACETYLASE-LIKE PROTEIN.
GN AT5G22650 OR MDJ22.7.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=98069011; PubMed=9405937;
RA Kotani H., Nakamura Y., Sato S., Kaneko T., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. II.
RT Sequence features of the regions of 1,044,062 bp covered by thirteen
RT physically assigned P1 clones.";
RL DNA Res. 4:291-300(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB006699; BAB11671.1; -.
DR EMBL; AY059893; AAL24375.1; -.
SQ SEQUENCE 306 AA; 32348 MW; A19274D43BDD142C CRC64;

```
  Query Match             99.3%;  Score 1578.5;  DB 10;  Length 306;
  Best Local Similarity   99.7%;  Pred. No. 9e-98;
  Matches  305;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQD 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQD 60

Qy      61 KFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 KFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTA 120

Qy     121 VTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMDVDE 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     121 VTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMDVDE 180

Qy     181 DDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQKT-E 239
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     181 DDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTEE 240

Qy     240 KKKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGGNNNKKPFNSGKQFGGSNNKGSNKG 299
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 KKKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGGNNNKKPFNSGKQFGGSNNKGSNKG 300

Qy     300 KGKGRA 305
           ||||||
Db     301 KGKGRA 306
```

RESULT 3
Q9FVE6
ID Q9FVE6 PRELIMINARY; PRT; 245 AA.
AC Q9FVE6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PUTATIVE HISTONE DEACETYLASE.
GN HD2A.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Wu K., Tian L., Malik K., Brown D., Miki B.;
RT "Functional analysis of HD2 histone deacetylase homologs in
RT Arabidopsis thaliana.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF195545; AAG28472.1; -.
DR InterPro; IPR000822; Znf-C2H2.
DR SMART; SM00355; ZnF_C2H2; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
KW DNA-binding; Metal-binding; Zinc-finger.
SQ SEQUENCE 245 AA; 26372 MW; E8E4F7E0A0D24F28 CRC64;

```
  Query Match             37.4%;  Score 594;  DB 10;  Length 245;
  Best Local Similarity   50.7%;  Pred. No. 2.4e-32;
  Matches  151;  Conservative   26;  Mismatches   63;  Indels   58;  Gaps   15;

Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASL-DCTVKSGESVVLSVTVGGAKLVIGTLSQ 59
           |||||: |       |||||  |:|:||||| :|  | || | | | ||   ||:||||
Db       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60

Qy      60 DKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPT 119
           :  ||:  ||||||||||||:  | :|:|:|||:||||   : | ::|:  | |||
Db      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGY-SEEEEEEEEEVPA---- 115

Qy     120 AVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMDVD 179
             |||  ||     ||||||||||| :   ||||||  :             ||  |
Db     116 -----GNAAKAV----AKPKAKPAEVKPAVD----DEEDESDSD----------GM--D 149

Qy     180 EDDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTE 239
           |||||    |||  |||| |||||   :||| ||:  | ||| |||| |     |||||:
Db     150 EDDSDG---EDS--EEEEPTPKKPAS-SKKRANETTPKAPVSAKKAKVAV----TPQKTD 199

Qy     240 KKKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGGNNNKKPFNSGKQFGGSNNKGSN 297
           :|           |||||:   |||||||  | | | : || ||||     |:||  :
Db     200 EK-----------KKGGKA---ANQSPKSASQVSCG-SCKKTPNSGNAL-ESHNKAKH 241
```

RESULT 4
O22238
ID O22238 PRELIMINARY; PRT; 257 AA.

```
AC   O22238;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE HISTONE DEACETYLASE.
GN   T32N15.8.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   de la Bastide M.R., Parnell L.D., Kaplan N., Gnoj L., Hameed A.,
RA   Schutz K., Hasegawa A., Gottesman T., Shohdy N., Granat S., Jensen K.,
RA   Johnson A.F., Lodhi M., Dedhia N., Martienssen R., McCombie W.R.;
RT   "A. thaliana BAC T32N15 from chromosome V.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC002534; AAB70032.1; -.
DR   InterPro; IPR000822; Znf-C2H2.
DR   Pfam; PF00096; zf-C2H2; 1.
DR   SMART; SM00355; ZnF_C2H2; 1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
KW   DNA-binding; Metal-binding; Zinc-finger.
SQ   SEQUENCE   257 AA;  27698 MW;  C44043876E6889D5 CRC64;
```

```
  Query Match             37.4%;  Score 594;  DB 10;  Length 257;
  Best Local Similarity   50.7%;  Pred. No. 2.5e-32;
  Matches  151;  Conservative   26;  Mismatches   63;  Indels   58;  Gaps   15;

Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASL-DCTVKSGESVVLSVTVGGAKLVIGTLSQ 59
           |||||: |       |||||  |:|:|||||| :|  | || | | | ||   ||:||||
Db       1 MEFWGIEVKSGKPVTVTPEEGILIHVSQASLGECKNKKGEFVPLHVKVGNQNLVLGTLST 60

Qy      60 DKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPT 119
           :  ||:  |||||||||||||:  | :|:|:|||:||||   : | ::|:  | |||
Db      61 ENIPQLFCDLVFDKEFELSHTWGKGSVYFVGYKTPNIEPQGY-SEEEEEEEEEVPA---- 115

Qy     120 AVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEKGMDVD 179
                |||  ||     |||||||||||| :    ||||||| :            ||  |
Db     116 -----GNAAKAV----AKPKAKPAEVKPAVD----DEEDESDSD-----------GM--D 149

Qy     180 EDDSDDDEEEDSEDEEEEEETPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTE 239
           |||||    |||  |||| |||||   :||| ||:  | ||| |||| |     |||||:
Db     150 EDDSDG---EDS--EEEEPTPKKPAS-SKKRANETTPKAPVSAKKAKVAV----TPQKTD 199

Qy     240 KKKGGHTATPHPAKKGGKSPVNANQSPKSGGQSSGGNNNKKPFNSGKQFGGSNNKGSN 297
           :|           |||||:   |||||||  | | | : || ||||     |:||  :
Db     200 EK-----------KKGGKA---ANQSPKSASQVSCG-SCKKTFNSGNAL-ESHNKAKH 241
```

```
RESULT   5
Q9M4T5
ID   Q9M4T5      PRELIMINARY;      PRT;   297 AA.
AC   Q9M4T5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE HISTONE DEACETYLASE HD2.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Dangl M., Brosch G., Haas H., Loidl P., Graessle S., Lusser A.;
RT   "Molecular characterization of type-2 histone deacetylases in higher
RT   plants.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF255711; AAF70196.1; -.
DR   InterPro; IPR000822; Znf-C2H2.
DR   SMART; SM00355; ZnF_C2H2; 1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
KW   DNA-binding; Metal-binding; Zinc-finger.
SQ   SEQUENCE   297 AA;  32502 MW;  EEB3AAC0D2C41C63 CRC64;
```

```
  Query Match             36.1%;  Score 574;  DB 10;  Length 297;
  Best Local Similarity   46.0%;  Pred. No. 6.2e-31;
  Matches  137;  Conservative   38;  Mismatches   93;  Indels   30;  Gaps   11;

Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQD 60
           |||||: | |    |  ||::  :|:|||:|  : |  :: |: |     |||||||| |
Db       1 MEFWGLEVKPGQTVKCEPEDERFLHLSQAALGESKKGSDNAVMYVKTDDQKLVIGTLSAD 60

Qy      61 KFPQISFDLVFDKEFELSHSGTKANVHFIGYK-SPNIEQD--DFTS---SDDEDVPEAVP 114
           ||||| ||||||||||||||:   |:| | ||| |   |:|  || |    |:|:  : :|
Db      61 KFPQIQFDLVFDKEFELSHTSKTASVFFSGYKVSQPAEEDEMDFDSEEVEDEEEEEKIIP 120

Qy     115 APAPTAVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEED--DDS 172
           ||      |||      |:     : |  :   :  | ::   |: ||:| :|:|  |:
Db     121 AP-----RANGK-----VEGKENEQKKQGKTDSSASKSKAAVNDDDDDDDSAEDDSQDED 170

Qy     173 EKGMDVDEDDSDDDEEEDSEDE-EEEETPKKPEPINKKRPNESVSKTPVSGKKAKPAAAP 231
             | |:| |:||  || ||| :||:||||||   |::  | | ||| | |||| |
Db     171 LSPEDDDDDSSEDDSSEDDEDESDEEDTPKKPE-TGKRKVAE1VLKTPSSDKKAK1A--- 226

Qy     232 ASTP--QKTEKKKGGHTATPHPAKKGGKSPVN---ANQSPKSGGQSSGGNNNKKPFNS 284
             ||  |||  ||| | |||||||:  |:|||     :||||||  |    :  | |||
Db     227 --TPSGQKTGDKKGVHVATPHPAKQASKTPVNDKSKEKSPKSGGGSISCKSCSKTFNS 282
```

```
RESULT   6
O24591
ID   O24591      PRELIMINARY;      PRT;   307 AA.
AC   O24591;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NUCLEOLAR HISTONE DEACETYLASE HD2-P39.
GN   HD2.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CUZCO 251;
RX   MEDLINE=97349336; PubMed=9204905;
RA   Lusser A., Brosch G., Loidl A., Haas H., Loidl P.;
RT   "Identification of maize histone deacetylase HD2 as an acidic
RT   nucleolar phosphoprotein.";
RL   Science 277:88-91(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CUZCO;
RA   Brosch G., Lusser A., Loidl P.;
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U82815; AAB63262.1; -.
DR   EMBL; AF026917; AAC61674.1; -.
DR   InterPro; IPR000822; Znf-C2H2.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
KW   DNA-binding; Zinc-finger.
SQ   SEQUENCE   307 AA;  33238 MW;  9CFF6B036D274FC1 CRC64;
```

```
  Query Match             34.2%;  Score 543.5;  DB 10;  Length 307;
  Best Local Similarity   44.8%;  Pred. No. 6.9e-29;
  Matches  130;  Conservative   45;  Mismatches   96;  Indels   19;  Gaps   11;
```

```
Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQD 60
           |||||: | | :  |  |    ::|:|||:|  : || :: :: | :   || ||||| |
Db       1 MEFWGLEVKPGSTVKCEPGYGFVLHLSQAALGESKKS-DNALMYVKIDDQKLAIGTLSVD 59

Qy      61 KFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTA 120
           | | | |||:|||||||||:    :| | |||     ::|    | ||  | :  |
Db      60 KNPHIQFDLIFDKEFELSHTSKTTSVFFTGYKVEQPFEEDEMDLDSEDEDEELNVP---V 116

Qy     121 VTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEED--DDSEKGMDV 178
           |  || |      |:  |  | |::  |  :| : |:: : |: |:|:||  |||::|:
Db     117 VKENGKADEKKQKSQEKAVAAPSKSSPDSKKSKDDDDSDEDETDDSDEDETDDSDEGLSS 176

Qy     179 DEDDSDDDEEEDSEDEEEEE--TPKKPEPINKKRPNE-SVSKTPVSGKKAKPAAAPASTP 235
           :| | |  :|:|: |:|||:  ||||||  : |||| | || |||:| |||| | |:|
Db     177 EEGDDDSSDEDDTSDDEEEDTPTPKKPE-VGKKRPAESSVLKTPLSDKKAK-VATPSS-- 232

Qy     236 QKTEKKKGG--HTATPHPAKKGGKSPVNANQSPKS--GGQSSGGNNNKKP 281
           |||  |||   | |||||||  ||: || ::| ||      |||:   ||
Db     233 QKTGGKKGAAVHVATPHPAK--GKTIVNNDKSVKSPKSAPKSGGSVPCKP 280


RESULT   7
Q9M4U5
ID   Q9M4U5      PRELIMINARY;      PRT;   303 AA.
AC   Q9M4U5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HISTONE DEACETYLASE 2 ISOFORM B.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Dangl M., Brosch G., Haas H., Loidl P., Lusser A.;
RT   "Molecular characterization of type-2 histone deacetylases in higher
RT   plants.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF254072; AAF68624.1; -.
DR   InterPro; IPR000822; Znf-C2H2.
DR   SMART; SM00355; ZnF_C2H2; 1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
KW   DNA-binding; Metal-binding; Zinc-finger.
SQ   SEQUENCE   303 AA;  32613 MW;  708E4627101BB67C CRC64;


  Query Match             34.0%;  Score 541;  DB 10;  Length 303;
  Best Local Similarity   44.1%;  Pred. No. 9.9e-29;
  Matches  135;  Conservative   36;  Mismatches   95;  Indels   40;  Gaps   12;

Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQD 60
           |||||: | | :  |  |    ::|:|||:|  : || :| :: | |   || ||||| |
Db       1 MEFWGLEVKPGSTVKCEPGHGFILHVSQAALGESKKS-DSALMYVKVDDKKLAIGTLSID 59

Qy      61 KFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTA 120
           |:||| |||||:|||||||:    :| | |||     : |    | ||  | :  |
Db      60 KYPQIQFDLVFNKEFELSHTSKTTSVFFSGYKVEQPIEGDEMDLDSEDEEEELNIP---V 116

Qy     121 VTANGNAGAAVVKADTKPKAKPAE--------------VKPAEEKPESDEEDESDDEDES 166
           :  ||       ||| | : |  |               | :::  :  :||||||  ||
Db     117 IKENG-------KADGKEEQKNQEKAVAATASKSSLGLEKKSKDDSDDSDEDESDDSDED 169

Qy     167 EEDDDSE-KGMDVDEDDSDDDEEEDSEDEEEEE--TPKKPEPINKKRPNESVSKTPVSGK 223
           : ||  | :|:  || | |  :|:|: |::|||  ||||||   |||  |:  |||:| |
Db     170 DSDDSDEGEGLSPDEGDDDSSDEDDTSDDDEEETPTPKKPE-AGKKRGAENALKTPLSDK 228

Qy     224 KAKPAAAPASTPQKTEKKKGG-HTATPHPAKKGGKSPVN----ANQSPKSGGQSSGGNNN 278
           |||  |  ||   |||  |||  | |||||||  ||:| |      :|||||| |    :
Db     229 KAKVATPPA---QKTGGKKGATHVATPHPAK--GKTPANNDKLTEKSPKSGG-SVPCKSC 282
```

```
Qy     279 KKPFNS 284
             | |||
Db     283 SKTFNS 288


RESULT   8
Q9M4U4
ID   Q9M4U4      PRELIMINARY;      PRT;   300 AA.
AC   Q9M4U4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HISTONE DEACETYLASE 2 ISOFORM C.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Dangl M., Brosch G., Haas H., Loidl P., Lusser A.;
RT   "Molecular characterization of type-2 histone deacetylases in higher
RT   plants.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF254073; AAF68625.1; -.
DR   InterPro; IPR000822; Znf-C2H2.
DR   SMART; SM00355; ZnF_C2H2; 1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
KW   DNA-binding; Zinc-finger.
SQ   SEQUENCE   300 AA;  32457 MW;  62F6C4148EB1E1FE CRC64;


  Query Match             33.5%;  Score 532;  DB 10;  Length 300;
  Best Local Similarity   44.0%;  Pred. No. 3.9e-28;
  Matches  124;  Conservative   44;  Mismatches   92;  Indels   22;  Gaps    9;

Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQD 60
           |||||: | | :  |  |    ::|:|||:|  : || :: :: | :   || ||||| |
Db       1 MEFWGLEVKPGSTVKCEPGYGFVLHLSQAALGESKKS-DNALMYVKIDDQKLAIGTLSVD 59

Qy      61 KFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAVPAPAPTA 120
           | | | |||:|||||||||:    :| | |||     ::|    | ||  | :  |
Db      60 KNPHIQFDLIFDKEFELSHTSKTTSVFFTGYKVEQPFEEDEMDLDSEDEDEELNVP---A 116

Qy     121 VTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEED--DDSEKGMDV 178
           |  || |      |:  |  | |::  |  :| : |:: : |: |:|:||  |||::|:
Db     117 VKENGKADEKKQKSQEKAVAAPSKSSPDSKKSKDDDDSDEDETDDSDEDETDDSDEGLSP 176

Qy     179 DEDDSDDDEEEDSEDEEEEE--TPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQ 236
           :| | |  :|:|: |:|||:  ||||||   |:    || |||:| |||| | |:|  |
Db     177 EEGDDDSSDEDDTSDDEEEDTPTPKKPEVGKKRAAESSVLKTPLSDKKAK-VATPSS--Q 233

Qy     237 KTEKKKGG--HTATPHPAKKGGKSPVNANQS-------PKSG 269
           ||  |||   | |||||||  ||: || ::|       ||||
Db     234 KTGGKKGAAVHVATPHPAK--GKTIVNNDKSVKSPKSAPKSG 273


RESULT   9
Q9M4T4
ID   Q9M4T4      PRELIMINARY;      PRT;   294 AA.
AC   Q9M4T4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE HISTONE DEACETYLASE HD2C (AT5G03740/F17C15_160).
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RA   Dangl M., Brosch G., Haas H., Loidl P., Graessle S., Lusser A.;
RT   "Molecular characterization of type-2 histone deacetylases in higher
RT   plants.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=COLUMBIA;
RX   MEDLINE=97471969; PubMed=9330910;
RA   Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
RA   Miyajima N., Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
RT   features of the 1.6 Mb regions covered by twenty physically assigned
RT   P1 clones.";
RL   DNA Res. 4:215-230(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Cheuk R., Chen H., Kim C.J., Meyers M.C., Shinn P., Banh J.,
RA   Bowser L., Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y.,
RA   Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA   Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M.,
RA   Palm C.J., Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M.,
RA   Seki M., Southwick A., Toriumi M., Yamada K., Yu G., Shinozaki K.,
RA   Davis R.W., Theologis A., Ecker J.R.;
RT   "Arabidopsis cDNA clones.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF255712; AAF70197.1; -.
DR   EMBL; AB005235; BAB08599.1; -.
DR   EMBL; AF372889; AAK49605.1; -.
DR   InterPro; IPR000822; Znf-C2H2.
DR   SMART; SM00355; ZnF_C2H2; 1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
KW   DNA-binding; Metal-binding; Zinc-finger.
SQ   SEQUENCE   294 AA;  31830 MW;  7B69F3BE8324C1D1 CRC64;


  Query Match             28.5%;  Score 453.5;  DB 10;  Length 294;
  Best Local Similarity   39.5%;  Pred. No. 6.4e-23;
  Matches  123;  Conservative   38;  Mismatches   85;  Indels   65;  Gaps   13;

Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGES-------VVLSVTVGGAKLV 53
           |||||| |       : |  | |||||| :|      |||       : | ||||  ||:
Db       1 MEFWGVEVKNGKPLHLDPGLDRLVHISQVAL------GESKNNVTEPIQLYVTVGSDKLL 54

Qy      54 IGTLSQDKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDEDVPEAV 113
           ||||| :||||:| ::| :: | |||:    :| | |||       | :  : ||:
Db      55 IGTLSHEKFPQLSTEIVLERNFALSHTWKNGSVFFSGYKV------DASDPEPEDL---- 104

Qy     114 PAPAPTAVTANGNAGAAVVKADTKPKAK------PAEVKPAEEKPESD--EEDESDDEDE 165
                 :    ||  ||  |  ||      | |   |::  ::|  ||| |||:|
Db     105 ---------IDDQLEAAGFKAAPKSAAKQVNFQLPNEDVKAKQDDDADGSEEDSSDDDDS 155

Qy     166 SEEDDDSEKGM----DVDEDDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKTPVS 221
               |: |: :    | :|||| ||||:||    |||||||||   |:    : || | |
Db     156 ENSGDEEEEKVTAESDSEEDDSSDDEEDDS---SEEETPKKPEEPKKRSAEPNSSKNPAS 212

Qy     222 GKKAKPAAAPASTPQKTEKKKGG-HTATPHPAKKGGKSPVNANQSPKSGGQSSGGNNNKK 280
           ||||       |||||: ||   | |||||:|: ||:          ||| |:|  : ::
Db     213 NKKAK-----FVTPQKTDSKKPHVHVATPHPSKQAGKN---------SGGGSTGETSKQQ 258

Qy     281 --PFNSGKQFG 289
             | ::|  ||
Db     259 QTPKSAG-AFG 268


RESULT  10
Q9LZR5
ID   Q9LZR5      PRELIMINARY;      PRT;   296 AA.
AC   Q9LZR5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HISTONE DEACETYLASE-LIKE PROTEIN.
GN   F17C15_160.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W., Rudd S.,
RA   Lemcke K., Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL162506; CAB82939.1; -.
DR   InterPro; IPR000822; Znf-C2H2.
DR   SMART; SM00355; ZnF_C2H2; 1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
KW   DNA-binding; Metal-binding; Zinc-finger.
SQ   SEQUENCE   296 AA;  32129 MW;  4B2DAABEAA1F1E9D CRC64;


  Query Match             28.4%;  Score 450.5;  DB 10;  Length 296;
  Best Local Similarity   38.3%;  Pred. No. 1e-22;
  Matches  118;  Conservative   39;  Mismatches   94;  Indels   57;  Gaps   12;

Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGES-------VVLSVTVGGAKLV 53
           |||||| |       : |  | |||||| :|      |||       : | ||||  ||:
Db       1 MEFWGVEVKNGKPLHLDPGLDRLVHISQVAL------GESKNNVTEPIQLYVTVGSDKLL 54

Qy      54 IGTLSQDKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDED----- 108
           ||||| :||||:| ::| :: | |||:    :| | |||  ::  |   ::  :
Db      55 IGTLSHEKFPQLSTEIVLERNFALSHTWKNGSVFFSGYKPEDLIDDQLEAAGFKGKWGLL 114

Qy     109 VPEAVPAPAPTAVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEE 168
           | | |  |   |         :   | | |         ::  :  ||| |||:|
Db     115 YPPAAPKSAAKQVNFQ------LPNEDVKAK--------QDDDADGSEEDSSDDDDSENS 160

Qy     169 DDDSEKGM----DVDEDDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKK 224
             |: |: :    | :|||| ||||:||    |||||||||   |:    : || | | ||
Db     161 GDEEEEKVTAESDSEEDDSSDDEEDDS---SEEETPKKPEEPKKRSAEPNSSKNPASNKK 217

Qy     225 AKPAAAPASTPQKTEKKKGG-HTATPHPAKKGGKSPVNANQSPKSGGQSSGGNNNKK--P 281
           ||       |||||: ||   | |||||:|: ||:          ||| |:|  : ::  |
Db     218 AK-----FVTPQKTDSKKPHVHVATPHPSKQAGKN---------SGGGSTGETSKQQQTP 263

Qy     282 FNSGKQFG 289
           ::|  ||
Db     264 KSAG-AFG 270


RESULT  11
Q94F81
ID   Q94F81      PRELIMINARY;      PRT;   286 AA.
AC   Q94F81;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HD2 TYPE HISTONE DEACETYLASE HDA106.
GN   HDA106.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. B73;
RA   Chandler V.L., Kaeppler S.M., Kaeppler H.F., Cone K.C.;
RT   "Sequences from the Plant Chromatin Consortium (NSF Plant Genome
RT   project 9975930).";
```

```
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. B73;
RA   Bergstrom D., Springer N.M., Schmitt L.T., Guthrie E., Sidorenko L.,
RA   Selinger D., Kaeppler S.M., Cone K.C.;
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF384033; AAK67143.1; -.
SQ   SEQUENCE   286 AA;  30808 MW;  E295E5CDFE4DEF1A CRC64;


  Query Match             22.9%;  Score 364;  DB 10;  Length 286;
  Best Local Similarity   33.1%;  Pred. No. 5.6e-17;
  Matches  106;  Conservative   43;  Mismatches   81;  Indels   90;  Gaps   13;

Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQD 60
           ||  |  | |          :  ::|:|||:|  : |: |: :||| :   |||:|||| :
Db       1 MEVGGQEVKPGATVSCKVGDGLVIHLSQAALGESKKASENAILSVNIDDKKLVLGTLSVE 60

Qy      61 KFPQISFDLVFDKEFELSHSGTKANVHFIGYKSP------NIEQDDFTSSDDEDVPEAVP 114
           | |||| ||||||:||| |:    :| | |||||      |  :|   ||||:|
Db      61 KHPQISCDLVFDKDFELPHNSKTRSVFFRGYKSPVPLFESNSGED---SSDEE------- 110

Qy     115 APAPTAVTANGNAGAAVVKADTKPKAKPAEVK-PAEEKPESDEEDE-------------S 160
                     :| |  |  :  |:|  | : |  |::|:                    |
Db     111 ----------------LKTDQIP-LQNNEIKISAAKVPAKDDDDDVFIILAMMMMIYSS 152

Qy     161 DDEDE-----------SEEDDDSEKGMDVDEDDSDDDEE--------EDSEDE---EEEE 198
           ||:|:           ||||| |::    :|||| |::|         :|| ||   |
Db     153 DDDDDDFTTSDSDNEMSEEDDSSDEDEMSEEDDSSDEDEMSGGADPSDDSSDESGSEHTS 212

Qy     199 TPKKPE-PINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTEKKKGGHTATPHPAK---- 253
            ||| :  : |||   ::      |||||          ::: :| |   :| ||||
Db     213 APKKTDVVVGKKR---AIKAEAPYGKKAK--------SEQSSQKTGDKASTSHPAKQSIK 261

Qy     254 ----KGGKSPVNANQSPKSG 269
               |  |:|    :|||||
Db     262 TPADKSRKTPTADKKSPKSG 281


RESULT  12
Q9M4T3
ID   Q9M4T3      PRELIMINARY;      PRT;   203 AA.
AC   Q9M4T3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   PUTATIVE HISTONE DEACETYLASE HD2D.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Dangl M., Brosch G., Haas H., Loidl P., Graessle S., Lusser A.;
RT   "Molecular characterization of type-2 histone deacetylases in higher
RT   plants.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF255713; AAF70198.1; -.
SQ   SEQUENCE   203 AA;  22651 MW;  69A1DA5941D7A293 CRC64;


  Query Match             16.5%;  Score 262.5;  DB 10;  Length 203;
  Best Local Similarity   30.9%;  Pred. No. 2.2e-10;
  Matches   80;  Conservative   33;  Mismatches   75;  Indels   71;  Gaps   11;

Qy       1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCT--VKSGESVVLSVTVGGAK--LVIGT 56
           |||||: : |    ||  ::  :|| || :|      ||  |:  : | :|  :   :||
Db       1 MEFWGIEIKPGKPFKVIQKDGFMVHASQVTLGDVEKVKKDETFAVYVKIGDDENGFMIGN 60

Qy      57 LSQDKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNI--EQDDFTSSDDEDVPEAVP 114
           ||| |||| | ||    |||:||: | ::|: |||:: :   | |:   || |
Db      61 LSQ-KFPQFSIDLYLGHEFEISHNST-SSVYLIGYRTFDAFDELDEEIDSDSE------- 111

Qy     115 APAPTAVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESEEDDDSEK 174
                                 |   : : | :   |  || |:| :||          |
Db     112 ------------------LDEYMEQQIAALPQNEINPEEDDESDSD----------EM 141

Qy     175 GMDVDEDDSDDDEEEDSEDEEEEETPKKPEPINKKRPNES--VSKTPVSGKKAKPAAAPA 232
           |:| |:| ||       |:: | | | |  | :|| || :  ::|    |||
Db     142 GLDEDDDSSD-------EEDVEAEAPLKVAPPSKKMPNGAFEIAK---GGKK-------- 183

Qy     233 STPQKTEKKKGGHTATPHP 251
                 |  ||    | |
Db     184 ------NKSSGGKKRCPFP 196


RESULT  13
Q26711
ID   Q26711      PRELIMINARY;      PRT;   312 AA.
AC   Q26711;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NOPP44/46.
OS   Trypanosoma brucei.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX   NCBI_TaxID=5691;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EATRO 164;
RX   MEDLINE=96279091; PubMed=8663171;
RA   Das A., Peterson G.C., Kanner S.B., Frevert U., Parsons M.;
RT   "A major tyrosine-phosphorylated protein of Trypanosoma brucei is a
RT   nucleolar RNA-binding protein.";
RL   J. Biol. Chem. 271:15675-15681(1996).
DR   EMBL; U53863; AAC47134.1; -.
SQ   SEQUENCE   312 AA;  33222 MW;  2AB8ED8D4416EB00 CRC64;


  Query Match             16.2%;  Score 258;  DB 5;  Length 312;
  Best Local Similarity   24.5%;  Pred. No. 7e-10;
  Matches   83;  Conservative   55;  Mismatches  117;  Indels   84;  Gaps   12;

Qy       3 FWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSGESVVLSVTVGGAKLVIGTLSQDK- 61
           |:|| |:      |    || :: ::| ::      || :: | ::  | :  | ||   :
Db       4 FYGVEVSAGQKVKPKIPEDHVLRLTQIAVPANA-SG-AITLVISFQGKEFTIATLDPKRS 61

Qy      62 FPQISFDLVFDKEFELSHSGT-KANVHFIGYKSPNIEQDDFTSSDDEDVPEA-------V 113
             |:  |||   |   : | |   :||  ||  |  :  ||   || |  :|       |
Db      62 LFQMGIDLVLTAEQGTTLSATGSGSVHLTGYVQPVDDGSDFDEEDDLDEVDAEDMESYGV 121

Qy     114 P-------APAPTAVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDES 166
           |       |  | :   :  |   |  | :   :   ::   |  :: :|||::||:||:
Db     122 PPKRGGKSAKFPVSEENSDEANGDGVDGDDESDDEEG-IEKDNEGIDAGDEDDNDDDDEA 180

Qy     167 EEDDDSEKGMDVDEDDSDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKKAK 226
            ::|||:     |:|| |||:::| :|::::|
Db     181 YDEDDSD-----DDDDDDDDDDDDDDDDDE-------------------------- 207

Qy     227 PAAAPASTPQKTEKK----------------KGGHTATPHPAKKGGKSPVNANQSPKSG 269
               |||  |:||::                 ||:        :||      |:    |
Db     208 ----PASKVQRTEQQGFRGGNRGGFRGGDRNGNGGNRGGFRGGDRGG------NRGNFRG 257

Qy     270 GQSSGGNNNKKPFNSGKQFGGSNN-----KGSNKGKGKG 303
           |  ||   |:  |  | : |   |      :| |:|  :|
Db     258 GDRSGNGGNRGNFRGGDRGGNRGNFRGGDRGGNRGNFRG 296


RESULT  14
O45181
ID   O45181      PRELIMINARY;      PRT;   798 AA.
AC   O45181;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
```

```
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 86.8 KDA PROTEIN.
GN   K07H8.10.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Fulton B., Hawkins J., Gattung S., Wohldmann P., Elliott G.;
RT   "The sequence of C. elegans cosmid K07H8.";
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF047659; AAC04430.1; -.
DR   InterPro; IPR002106; AA_tRNA_ligase_II.
DR   InterPro; IPR000504; RRM.
DR   Pfam; PF00076; rrm; 2.
DR   SMART; SM00360; RRM; 2.
DR   PROSITE; PS00179; AA_TRNA_LIGASE_II_1; UNKNOWN_1.
DR   PROSITE; PS50102; RRM; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   798 AA;  86819 MW;  6209063CD13FAB26 CRC64;


  Query Match             14.5%;  Score 230;  DB 5;  Length 798;
  Best Local Similarity   31.7%;  Pred. No. 1.4e-07;
  Matches   72;  Conservative   32;  Mismatches   83;  Indels   40;  Gaps    9;

Qy      48 GGAKLVIGTLSQDKFPQISFDLVFDKEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDE 107
           ||:::       ||  |   |   |:| |:          |     | :| |  :  |:|
Db     181 GGSRVGF----SDKRKQYDSDGDDDEEEEVPKKKLATGTPFAKQTKP-VEDD--SGEDEE 233

Qy     108 DVPEAVPAPAPTAVTANGNAGAAVVKADTKPKAKPAEVKPAEEKPESDEEDESDDEDESE 167
           |  |:   | |                    |  || |    ||   :||:::| |||:| |
Db     234 DEEESDEEPQP----------------PKKTAKSAIATAAE---DSDDDEEDDDEEEEE 273

Qy     168 EDDDS---EKGMDVDEDDSDDDEEE----DSEDEEEEETPKKPEPINKKRPNESVSKTPV 220
           |||:   :| : |:||: |:::||    | |||||||  :: ||: |    :|  |:
Db     274 SDDDAPQVKKSVAVEEDEEDEEDEEEIEMDEEDEEEEEDEEETEPVAKTPAVKSSLKSVD 333

Qy     221 SGKKAKPAAAPASTPQKTEKKKGGHTATPHPAKKGGKSPVNANQSPK 267
           | |  :   :  :||   :||     |||||||    :   |   ||
Db     334 STKGKQVNLSKVATPTTVDKKA---PATPHPAKTSKNA---AGDVPK 374


RESULT  15
Q9NLA3
ID   Q9NLA3      PRELIMINARY;      PRT;   346 AA.
AC   Q9NLA3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NUCLEIC ACID-ASSOCIATED PROTEIN 36.
OS   Asterina pectinifera (Starfish).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
OC   Asteroidea; Valvatacea; Valvatida; Asterinidae; Asterina.
OX   NCBI_TaxID=7594;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20098505; PubMed=10632699;
RA   Nakajima H., Matoba K., Matsumoto Y., Hongo T., Kiritaka K.,
RA   Sugino H., Nagamatsu Y., Hamaguchi Y., Ikegami S.;
RT   "Molecular characterization of a novel nucleolar protein in starfish
RT   oocytes which is phosphorylated before and during oocyte maturation.";
RL   Eur. J. Biochem. 267:295-304(2000).
DR   EMBL; AB003034; BAA90827.1; -.
DR   InterPro; IPR004301; Nucleoplasmin.
DR   Pfam; PF03066; Nucleoplasmin; 1.
SQ   SEQUENCE   346 AA;  39005 MW;  BA88FC06167E2FDC CRC64;


  Query Match             13.2%;  Score 210.5;  DB 5;  Length 346;
  Best Local Similarity   26.8%;  Pred. No. 1.1e-06;
  Matches   91;  Conservative   34;  Mismatches  123;  Indels   91;  Gaps   15;

Qy       3 FWGVAVT-PKNATKVTP----------EEDSLVH---ISQASLDCTVKSGESVVLSVTV 47
           ||| ::|  |   |  |           : | : |   : || |    | ||  |:
Db       6 FWGDSLTGTKKEVKWNPSLDDEDDFDNLDSDGIQHFLFLKQAVLGANAKEGERNVVE--- 62

Qy      48 GGAKLVIGTLSQD----KFPQISFDLVFDKEFELSHSGTKANVHFI----------GYK 92
                 | | : |    | |  |  |  ::   |   | :  | ||           |
Db      63 ------IETENFDGDNVKQPLFSLKLGLNESSPLD-IGIQPPVTFILTAGSGPVFLSGQH 115

Qy      93 SPNIEQDDFTSSDDEDVPEAVPAPAPTAVTANGNAGAAVVKADTKP---------KAKPA 143
             |  ||    :::|  |          |  | :    | |:  :|         | | |
Db     116 MIEISADDEEELEEDDEEE----EEEDEVEVNASPDLPVAKSKKRPLSTSDGTAKKTKMA 171

Qy     144 EVKPAEEKPESDEEDESD----------DEDESEEDD----------DSEKGMDVDEDD 182
           ::    :| | |:|:| |          |:|| :|||           : |:  | ||||
Db     172 KLDKDADKKEDDDEEEDDEEEDEVMAMMDDDEDDEDDEDFEGGEDDEEEDEEESDEDEDD 231

Qy     183 SDDDEEEDSEDEEEEETPKKPEPINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTEKKK 242
           ||:|||: || |:||:|:||             ||   |||  :      :|  | : |
Db     232 EDDNEEEEEED-EDEESPEKP----------LKTTAKGKKGQ---MNGTTTAKGDNKP 275

Qy     243 GGHTATPHPAKKG-GKSPVNANQSPKSGGQSSGGNNNKK 280
             |     ||  |  | |    |   | |     |:
Db     276 KAKAKTDTKLVKGKAKKKVLALDEIKGKLQESSNVPKKE 314


Search completed: May  2, 2002, 22:30:24
Job time: 1585 sec
```